Mineharu Nakayama (Ed.)
**Handbook of Japanese Psycholinguistics**

# Handbooks of Japanese Language and Linguistics

Edited by
Masayoshi Shibatani
Taro Kageyama

## Volume 9

# Handbook of Japanese Psycholinguistics

Edited by
Mineharu Nakayama

DE GRUYTER
MOUTON

ISBN 978-1-61451-165-6
e-ISBN (PDF) 978-1-61451-121-2
e-ISBN (EPUB) 978-1-5015-0058-9
ISSN 2199-2851

**Library of Congress Cataloging-in-Publication Data**
A CIP catalog record for this book has been applied for at the Library of Congress.

**Bibliographic information published by the Deutsche Nationalbibliothek**
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.


Cover images: Jonathan Kashiwagi-Wood, Seth Goss, Jennifer Nakayama
Typesetting: RoyalStandard, Hong Kong
Printing and binding: CPI books GmbH, Leck
♾ Printed on acid-free paper
Printed in Germany

www.degruyter.com

# Preface

The project of compiling a series of comprehensive handbooks covering major fields of Japanese linguistics started in 2011, when Masayoshi Shibatani received a commission to edit such volumes as series editor from De Gruyter Mouton. As the planning progressed, with the volume titles selected and the volume editors assigned, the enormity of the task demanded the addition of a series co-editor. Taro Kageyama, Director-General of the National Institute for Japanese Language and Linguistics (NINJAL), was invited to join the project as a series co-editor. His participation in the project opened the way to make it a joint venture between NINJAL and De Gruyter Mouton. We are pleased to present the *Handbooks of Japanese Language and Linguistics* (*HJLL*) as the first materialization of the agreement of academic cooperation concluded between NINJAL and De Gruyter Mouton.

The HJLL Series is composed of twelve volumes, primarily focusing on Japanese but including volumes on the Ryukyuan and Ainu languages, which are also spoken in Japan, as well as some chapters on Japanese Sign Language in the applied linguistics volume.

- Volume 1: *Handbook of Japanese Historical Linguistics*
- Volume 2: *Handbook of Japanese Phonetics and Phonology*
- Volume 3: *Handbook of Japanese Lexicon and Word Formation*
- Volume 4: *Handbook of Japanese Syntax*
- Volume 5: *Handbook of Japanese Semantics and Pragmatics*
- Volume 6: *Handbook of Japanese Contrastive Linguistics*
- Volume 7: *Handbook of Japanese Dialects*
- Volume 8: *Handbook of Japanese Sociolinguistics*
- Volume 9: *Handbook of Japanese Psycholinguistics*
- Volume 10: *Handbook of Japanese Applied Linguistics*
- Volume 11: *Handbook of the Ryukyuan Languages*
- Volume 12: *Handbook of the Ainu Language*

Surpassing all currently available reference works on Japanese in both scope and depth, the *HJLL* series provides a comprehensive survey of nearly the entire field of Japanese linguistics. Each volume includes a balanced selection of articles contributed by established linguists from Japan as well as from outside Japan and is critically edited by volume editors who are leading researchers in their individual fields. Each article reviews milestone achievements in the field, provides an overview of the state of the art, and points to future directions of research. The twelve titles are thus expected individually and collectively to contribute not only to the enhancement of studies on Japanese on the global level but also to the opening up of new perspectives for general linguistic research from both empirical and theoretical standpoints.

The *HJLL* project has been made possible by the active and substantial participation of numerous people including the volume editors and authors of individual

chapters. We would like to acknowledge with gratitude the generous support, both financial and logistic, given to this project by NINJAL. We are also grateful to John Haig (retired professor of Japanese linguistics, the University of Hawai'i at Mānoa), serving as copy-editor for the series. In the future, more publications are expected to ensue from the NINJAL-Mouton academic cooperation.

Masayoshi Shibatani, Deedee McMurtry Professor of Humanities and Professor of Linguistics, Rice University/Professor Emeritus, Kobe University
Taro Kageyama, Director-General, National Institute for Japanese Language and Linguistics (NINJAL)/Professor Emeritus, Kwansei Gakuin University

Masayoshi Shibatani and Taro Kageyama

# Introduction to the *Handbooks of Japanese Language and Linguistics*

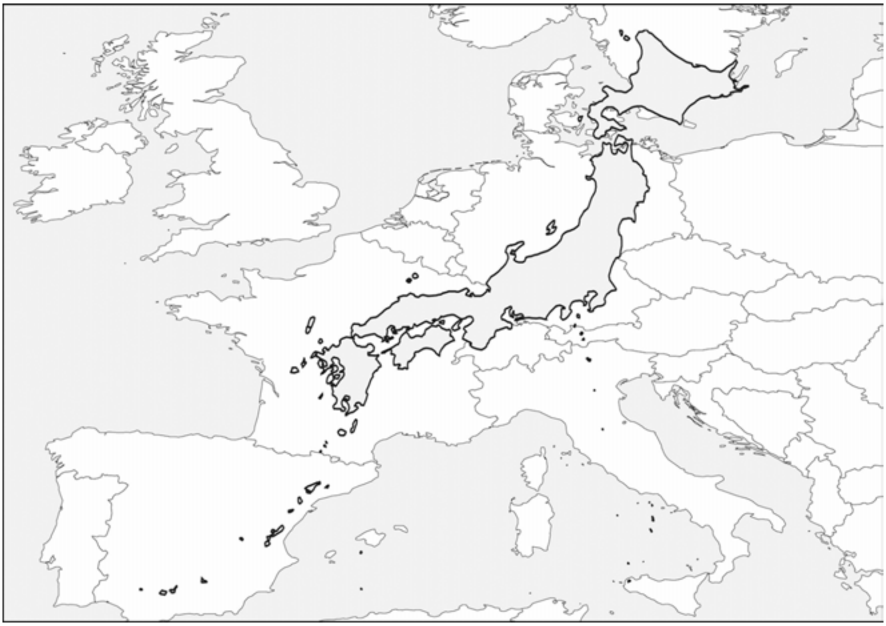
Comprising twelve substantial volumes, the *Handbooks of Japanese Language and Linguistics* (HJLL) series provides a comprehensive survey of practically all the major research areas of Japanese linguistics on an unprecedented scale, together with surveys of the endangered languages spoken in Japan, Ryukyuan and Ainu. What follows are introductions to the individual handbooks, to the general conventions adopted in this series, and the minimum essentials of contemporary Standard Japanese. Fuller descriptions of the languages of Japan, Japanese grammar, and the history of the Japanese language are available in such general references as Martin (1975), Shibatani (1990), and Frellesvig (2010).

## 1 Geography, Population, and Languages of Japan

Japan is situated in the most populous region of the world – Asia, where roughly one half of the world population of seven billion speak a variety of languages, many of which occupy the top tier of the ranking of the native-speaker population numbers. Japanese is spoken by more than 128 million people (as of 2013), who live mostly in Japan but also in Japanese emigrant communities around the world, most notably Hawaii, Brazil and Peru. In terms of the number of native speakers, Japanese ranks ninth among the world's languages. Due partly to its rich and long literary history, Japanese is one of the most intensely studied languages in the world and has received scrutiny under both the domestic grammatical tradition and those developed outside Japan such as the Chinese philological tradition, European structural linguistics, and generative grammar developed in America. The *Handbooks of Japanese Language and Linguistics* intend to capture the achievements garnered over the years through analyses of a wide variety of phenomena in a variety of theoretical frameworks.

As seen in Map 1, where Japan is shown graphically superimposed on Continental Europe, the Japanese archipelago has a vast latitudinal extension of approximately 3,000 kilometers ranging from the northernmost island, roughly corresponding to Stockholm, Sweden, to the southernmost island, roughly corresponding to Sevilla, Spain.

Contrary to popular assumption, Japanese is not the only language native to Japan. The northernmost and southernmost areas of the Japanese archipelago are inhabited by people whose native languages are arguably distinct from Japanese. The southernmost sea area in Okinawa Prefecture is dotted with numerous small islands



**Map 1:** *Japan as overlaid on Europe*
Source: Shinji Sanada. 2007. *Hōgen wa kimochi o tsutaeru [Dialects convey your heart].* Tokyo: Iwanami, p. 68.

where Ryukyuan languages are spoken. Until recent years, Japanese scholars tended to treat Ryukyuan language groups as dialects of Japanese based on fairly transparent correspondences in sounds and grammatical categories between mainland Japanese and Ryukyuan, although the two languages are mutually unintelligible. Another reason that Ryukyuan languages have been treated as Japanese dialects is that Ryukyuan islands and Japan form a single nation. In terms of nationhood, however, Ryukyu was an independent kingdom until the beginning of the seventeenth century, when it was forcibly annexed to the feudal domain of Satsuma in southern Kyushu.

A more recent trend is to treat Ryukyuan as forming a branch of its own with the status of a sister language to Japanese, following the earlier proposals by Chamberlain (1895) and Miller (1971). Many scholars specializing in Ryukyuan today even confer language status to different language groups within Ryukyuan, such as Amami language, Okinawan language, Miyako language, etc., which are grammatically distinct to the extent of making them mutually unintelligible. The prevailing view now has Japanese and Ryukyuan forming the Japonic family as daughter languages of Proto-Japonic. HJLL follows this recent trend of recognizing Ryukyuan as a sister language to Japanese and devotes one full volume to it. The ***Handbook of the Ryukyuan Languages*** provides the most up-to-date answers pertaining to Ryukyuan

language structures and use, and the ways in which these languages relate to Ryukyuan society and history. Like all the other handbooks in the series, each chapter delineates the boundaries and the research history of the field it addresses, comprises the most important and representative information on the state of research, and spells out future research desiderata. This volume also includes a comprehensive bibliography of Ryukyuan linguistics.

The situation with Ainu, another language indigenous to Japan, is much less clear as far as its genealogy goes. Various suggestions have been made relating Ainu to Paleo-Asiatic, Ural-Altaic, and Malayo-Polynesian or to such individual languages as Gilyak and Eskimo, besides the obvious candidate of Japanese as its sister language. The general consensus, however, points to the view that Ainu is related to Japanese quite indirectly, if at all, via the Altaic family with its Japanese-Korean subbranch (see Miller 1971; Shibatani 1990: 5–7 for an overview). Because Ainu has had northern Japan as its homeland and because HJLL is also concerned with various aspects of Japanese linguistics scholarship in general, we have decided to include a volume devoted to Ainu in this series. The ***Handbook of the Ainu Language*** outlines the history and current state of the Ainu language, offers a comprehensive survey of Ainu linguistics, describes major Ainu dialects in Hokkaido and Sakhalin, and devotes a full section to studies dealing with typological characteristics of the Ainu language such as polysynthesis and incorporation, person marking, plural verb forms, and aspect and evidentials.

## 2 History

Japan's rich and long literary history dates back to the seventh century, when the Japanese learned to use Chinese characters in writing Japanese. Because of the availability of abundant philological materials, the history of the Japanese language has been one of the most intensely pursued fields in Japanese linguistics. While several different divisions of Japanese language history have been proposed, Frellesvig (2010) proposes the following four linguistic periods, each embracing the main political epochs in Japanese history.

| | | | |
|---|---|---|---|
| 1. | Old Japanese | 700–800 | (Nara period, 712–794) |
| 2. | Early Middle Japanese | 800–1200 | (Heian period, 794–1185) |
| 3. | Late Middle Japanese | 1200–1600 | (Kamakura period, 1185–1333; Muromachi period, 1333–1573) |
| 4. | Modern Japanese | 1600– | (Edo, 1603–1868; Meiji, 1868–1912; Taishō, 1912–1926; Shōwa, 1926–1989; Heisei, 1989–) |

This division reflects a major gulf between Pre-modern and Modern Japanese caused by some radical changes in linguistic structure during the Late Middle Japanese period. Modern Japanese is often further subdivided into Early Modern (Edo, 1603–1868), Modern (Meiji, 1868–1912; Taishō, 1912–1926), and Present-day Japanese (Shōwa, 1926–1989; Heisei, 1989–).

The ***Handbook of Japanese Historical Linguistics*** will present the latest research on better studied topics, such as segmental phonology, accent, morphology, and some salient syntactic phenomena such as focus constructions. It will also introduce areas of study that have traditionally been underrepresented, ranging from syntax and Sinico-Japanese (*kanbun*) materials to historical pragmatics, and demonstrate how they contribute to a fuller understanding of the overall history of Japanese, as well as outlining larger-scale tendencies and directions in changes that have taken place within the language over its attested history. Major issues in the reconstruction of prehistoric Japanese and in the individual historical periods from Old Japanese to Modern Japanese are discussed including writing and the materials for historical studies, influences of Sinico-Japanese on Japanese, the histories of different vocabulary strata, the history of honorifics and polite language, generative diachronic syntax, and the development of case marking.

## 3 Geographic and Social Variations

Because of the wide geographical spread of the Japanese archipelago from north to south, characterized by high mountain ranges, deep valleys, and wide rivers as well as numerous islands, Japanese has developed a multitude of dialects, many of which differ from each other in a way more or less like current descendants of the Romance language family. Like the historical studies, the research tradition of dialect studies has a unique place in Japanese linguistics, which has also attracted a large number of students, amateur collectors of dialect forms as well as professional linguists. The ***Handbook of Japanese Dialects*** surveys the historical backdrop of the theoretical frameworks of contemporary studies in Japanese geolinguistics and includes analyses of prominent research topics in cross-dialectal perspectives, such as accentual systems, honorifics, verbs of giving, and nominalizations. The volume also devotes large space to sketch grammars of dialects from the northern island of Hokkaido to the southern island of Kyushu, allowing a panoramic view of the differences and similarities in the representative dialects throughout Japan.

Besides the physical setting fostering geographic variations, Japanese society has experienced several types of social structure over the years, starting from the time of the nobility and court life of the Old and Early Middle Japanese periods, through the caste structure of the feudalistic Late Middle and Early Modern Japanese periods, to the modern democratic society in the Modern and Present-day Japanese

periods. These different social structures spawned a variety of social dialects including power- and gender-based varieties of Japanese. The ***Handbook of Japanese Sociolinguistics*** examines a wide array of sociolinguistic topics ranging from the history of Japanese sociolinguistics, including foreign influences and internal innovations, to the central topics of variations due to social stratification, gender differences, and discourse genre. Specific topics include honorifics and women's speech, critical discourse analysis, pragmatics of political discourse, contact-induced change, emerging new dialects, Japanese language varieties outside Japan, and language policy.

# 4 Lexicon and Phonology

The literary history of Japan began with early contacts with China. Chinese apparently began to enrich the Japanese lexicon in even pre-historic periods, when such deeply assimilated words as *uma* 'horse' and *ume* 'plum' are believed to have entered the language. Starting in the middle of the sixth century, when Buddhism reached Japan, Chinese, at different periods and from different dialect regions, has continuously contributed to Japanese in an immeasurable way affecting all aspects of grammar, but most notably the lexicon and the phonological structure, which have sustained further and continuous influences from European languages from the late Edo period on. Through these foreign contacts, Japanese has developed a complex vocabulary system that is composed of four lexical strata, each with unique lexical, phonological, and grammatical properties: native Japanese, mimetic, Sino-Japanese, and foreign (especially English).

The ***Handbook of Japanese Lexicon and Word Formation*** presents a comprehensive survey of the Japanese lexicon, word formation processes, and other lexical matters seen in the four lexical strata of contemporary Japanese. The agglutinative character of the language, coupled with the intricate system of vocabulary strata, makes it possible for compounding, derivation, conversion, and inflection to be closely intertwined with syntactic structure, giving rise to theoretically intriguing interactions of word formation processes and syntax that are not easily found in inflectional, isolating, or polysynthetic types of languages. The theoretically oriented studies associated with these topics are complemented by those oriented toward lexical semantics, which also bring to light theoretically challenging issues involving the morphology-syntax interface.

The four lexical strata characterizing the Japanese lexicon are also relevant to Japanese phonology as each stratum has some characteristic sounds and sound combinations not seen in the other strata. The ***Handbook of Japanese Phonetics and Phonology*** describes and analyzes the basic phonetic and phonological structures of modern Japanese with main focus on standard Tokyo Japanese, relegating the topics of dialect phonetics and phonology to the *Handbook of Japanese Dialects*.

The handbook includes several chapters dealing with phonological processes unique to the Sino-Japanese and foreign strata as well as to the mimetic stratum. Other topics include word tone/accent, mora-timing, sequential voicing (*rendaku*), consonant geminates, vowel devoicing and diphthongs, and the appearance of new consonant phonemes. Also discussed are phonetic and phonological processes within and beyond the word such as rhythm, intonation, and the syntax-phonology interface, as well as issues bearing on other subfields of linguistics such as historical and corpus linguistics, L1 phonology, and L2 research.

## 5 Syntax and Semantics

Chinese loans have also affected Japanese syntax, though the extent is unclear to which they affected Japanese semantics beyond the level of lexical semantics. In particular, Chinese loans form two distinct lexical categories in Japanese – verbal nouns, forming a subcategory of the noun class, and adjectival nouns (*keiyō dōshi*), which are treated as forming major lexical categories, along with noun, verb, and adjective classes, by those who recognize this as an independent category. The former denote verbal actions, and, unlike regular nouns denoting objects and thing-like entities, they can function as verbs by combining with the light verb *suru*, which is obviously related to the verb *suru* 'do'. The nominal-verbal Janus character of verbal nouns results in two widely observed syntactic patterns that are virtually synonymous in meaning; e.g., *benkyoo-suru* (studying-DO) 'to study' and *benkyoo o suru* (studying ACC do) 'do studying'. As described in the *Handbook of Japanese Lexicon and Word Formation*, the lexical category of adjectival noun has been a perennial problem in the analysis of Japanese parts of speech. The property-concept words, e.g., *kirei* 'pretty', *kenkoo* 'health/healthy', falling in this class do not inflect by themselves unlike native Japanese adjectives and, like nouns, require the inflecting copula *da* in the predication function – hence the label of adjectival noun for this class. However, many of them cannot head noun phrases – the hallmark of the nominal class – and some of them even yield nouns via *-sa* nominalization, which is not possible with regular nouns.

The Lexicon-Word Formation handbook and the ***Handbook of Japanese Syntax*** make up twin volumes because many chapters in the former deal with syntactic phenomena, as the brief discussion above on the two Sino-Japanese lexical categories clearly indicates. The syntax handbook covers a vast landscape of Japanese syntax from three theoretical perspectives: (1) traditional Japanese grammar, known as *kokugogaku* (lit. national-language study), (2) the functional approach, and (3) the generative grammar framework. Broad issues analyzed include sentence types and their interactions with grammatical verbal categories, grammatical relations (topic, subject, etc.), transitivity, nominalization, grammaticalization, voice (passives and

causatives), word order (subject, scrambling, numeral quantifier, configurationality), case marking (*ga/no* conversion, morphology and syntax), modification (adjectives, relative clause), and structure and interpretation (modality, negation, prosody, ellipsis). These topics have been pursued vigorously over many years under different theoretical persuasions and have had important roles in the development of general linguistic theory. For example, the long sustained studies on the grammatical of subject and topic in Japanese have had significant impacts on the study of grammatical relations in European as well as Austronesian languages. In the study of word order, the analysis of Japanese numeral quantifiers is used as one of the leading pieces of evidence for the existence of a movement rule in human language. Under case marking, the way subjects are case-marked in Japanese has played a central role in the study of case marking in the Altaic language family. Recent studies of nominalizations have been central to the analysis of their modification and referential functions in a wide variety of languages from around the globe with far-reaching implications to past studies of such phenomena as parts of speech, (numeral) classifiers, and relative clauses. And the study of how in Japanese prosody plays a crucial role in interpretation has become the basis of some important recent developments in the study of wh-questions.

The ***Handbook of Japanese Semantics and Pragmatics*** presents a collection of studies on linguistic meaning in Japanese, either as conventionally encoded in linguistic form (the field of semantics) or as generated by the interaction of form with context (the field of pragmatics). The studies are organized around a model that has long currency in traditional Japanese grammar, whereby the linguistic clause consists of a multiply nested structure centered in a propositional core of objective meaning around which forms are deployed that express progressively more subjective meaning as one moves away from the core toward the periphery of the clause. Following this model, the topics treated in this volume range from aspects of meaning associated with the propositional core, including elements of meaning structured in lexical units (lexical semantics), all the way to aspects of meaning that are highly subjective, being most grounded in the context of the speaker. In between these two poles of the semantics-pragmatics continuum are elements of meaning that are defined at the level of propositions as a whole or between different propositions (propositional logic) and forms that situate propositions in time as events and those situating events in reality including non-actual worlds, e.g., those hoped for (desiderative meaning), denied (negation), hypothesized (conditional meaning), or viewed as ethically or epistemologically possible or necessary (epistemic and deontic modality). Located yet closer to the periphery of the Japanese clause are a rich array of devices for marking propositions according to the degree to which the speaker is committed to their veracity, including means that mark differing perceptual and cognitive modalities and those for distinguishing information variously presupposed.

These studies in Japanese syntax and semantics are augmented by cross-linguistic studies that examine various topics in these fields from the perspectives of language

universals and the comparative study of Japanese and another language. The ***Handbook of Japanese Contrastive Linguistics*** sets as its primary goal uncovering principled similarities and differences between Japanese and other languages around the globe and thereby shedding new light on the universal and language-particular properties of Japanese. Topics ranging from inalienable possession to numeral classifiers, from spatial deixis to motion typology, and from nominalization to subordination, as well as topics closely related to these phenomena are studied in the typological universals framework. Then various aspects of Japanese such as resultative-progressive polysemy, entailment of event realization, internal-state predicates, topic constructions, and interrogative pronouns, are compared and contrasted with individual languages including Ainu, Koryak, Chinese, Korean, Newar, Thai, Burmese, Tagalog, Kapampangan, Lamaholot, Romanian, French, Spanish, German, English, Swahili, Sidaama, and Mayan languages.

# 6 Psycholinguistics and Applied Linguistics

HJLL includes two volumes containing topics related to wider application of Japanese linguistics and to those endeavors seeking grammar-external evidence for the psycho-neurological reality of the structure and organization of grammar. By incorporating the recent progress in the study of the cognitive processes and brain mechanisms underlying language use, language acquisition, and language disorder, the ***Handbook of Japanese Psycholinguistics*** discusses the mechanisms of language acquisition and language processing. In particular, the volume seeks answers to the question of how Japanese is learned/acquired as a first or second language, and pursues the question of how we comprehend and produce Japanese sentences. The chapters in the acquisition section allow readers to acquaint themselves with issues pertaining to the question of how grammatical features (including pragmatic and discourse features) are acquired and how our brain develops in the language domain, with respect to both language-particular and universal features. Specific topics dealt with include Japanese children's perceptual development, the conceptual and grammatical development of nouns, Japanese specific language impairment, narrative development in the L1 cognitive system, L2 Japanese acquisition and its relation to L1 acquisition. The language processing section focuses on both L1 and L2 Japanese processing and covers topics such as the role of prosodic information in production/comprehension, the processing of complex grammatical structures such as relative clauses, the processing issues related to variable word order, and lexical and sentence processing in L2 by speakers of a different native language.

The ***Handbook of Japanese Applied Linguistics*** complements the Psycholinguistics volume by examining language acquisition from broader sociocultural per-

spectives, i.e., language as a means of communication and social behavioral system, emphasizing pragmatic development as central to both L1 and L2 acquisition and overall language/human development. Topics approached from these perspectives include the role of caregiver's speech in early language development, literacy acquisition, and acquisition of writing skills. Closely related to L1 and L2 acquisition/development are studies of bilingualism/multilingualism and the teaching and learning of foreign languages, including Japanese as a second language, where topics discussed include cross-lingual transfer from L1 to L2, learning errors, and proficiency assessment of second language acquisition. Chapters dealing with topics more squarely falling in the domain of applied linguistics cover the issues in corpus/computational linguistics (including discussions on CHILDES for Japanese and the KY corpus widely-used in research on Japanese as a second language), clinical linguistics (including discussions on language development in children with hearing impairment and other language disorders, with Down syndrome, or autism), and translation and interpretation. Technically speaking, Japanese Sign Language is not a variety of Japanese. However, in view of the importance of this language in Japanese society and because of the rapid progress in sign language research in Japan and abroad and what it has to offer to the general theory of language, chapters dealing with Japanese Sign Language are also included in this volume.

# 7 Grammatical Sketch of Standard Japanese

The following pages offer a brief overview of Japanese grammar as an aid for a quick grasp of the structure of Japanese that may prove useful in studying individual, thematically organized handbooks of this series. One of the difficult problems in presenting non-European language materials using familiar technical terms derived from the European grammatical tradition concerns mismatches between what the glosses may imply and what grammatical categories they are used to denote in the description. We will try to illustrate this problem below as a way of warning not to take all the glosses at their face value. But first some remarks are in order about the conventions of transcription of Japanese, glossing of examples, and their translations used in this series.

## 7.1 Writing, alphabetic transcription, and pronunciation

Customarily, Japanese is written by using a mixture of Chinese characters (for content words), *hiragana* (for function words such as particles, suffixes and inflectional endings), *katakana* (for foreign loans and mimetics), and sometimes Roman alphabet.

Because Japanese had no indigenous writing system, it developed two phonogram systems of representing a phonological unit of "mora", namely *hiragana* and *katakana*, by simplifying or abbreviating (parts of) Chinese characters. *Hiragana* and *katakana* syllabaries are shown in Table 1, together with the alphabetic transcriptions adopted in the HJLL series.

**Table 1:** *Alphabetic transcriptions adopted in HJLL*

| **transcription** | *a* | *ka* | *sa* | *ta* | *na* | *ha* | *ma* | *ya* | *ra* | *wa* | *n* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *hiragana* | あ | か | さ | た | な | は | ま | や | ら | わ | ん |
| *katakana* | ア | カ | サ | タ | ナ | ハ | マ | ヤ | ラ | ワ | ン |
| **transcription** | *i* | *ki* | *si* | *ti* | *ni* | *hi* | *mi* | – | *ri* | – | |
| *hiragana* | い | き | し | ち | に | ひ | み | – | り | – | |
| *katakana* | イ | キ | シ | チ | ニ | ヒ | ミ | – | リ | – | |
| **transcription** | *u* | *ku* | *su* | *tu* | *nu* | *hu* | *mu* | *yu* | *ru* | – | |
| *hiragana* | う | く | す | つ | ぬ | ふ | む | ゆ | る | – | |
| *katakana* | ウ | ク | ス | ツ | ヌ | フ | ム | ユ | ル | – | |
| **transcription** | *e* | *ke* | *se* | *te* | *ne* | *he* | *me* | – | *re* | – | |
| *hiragana* | え | け | せ | て | ね | へ | め | – | れ | – | |
| *katakana* | エ | ケ | セ | テ | ネ | ヘ | メ | – | レ | – | |
| **transcription** | *o* | *ko* | *so* | *to* | *no* | *ho* | *mo* | *yo* | *ro* | *o* | |
| *hiragana* | お | こ | そ | と | の | ほ | も | よ | ろ | を | |
| *katakana* | オ | コ | ソ | ト | ノ | ホ | モ | ヨ | ロ | ヲ | |

Because of phonological change, the columns indicated by strikethroughs have no letters in contemporary Japanese, although they were filled in with special letters in classical Japanese. If all the strikethroughs were filled, the chart will contain 50 letters for each of *hiragana* and *katakana*, so the syllabary chart is traditionally called *Gojū-on zu* (chart of 50 sounds). To these should be added the letter ん or ン representing a moraic nasal [N], on the rightmost column.

The "50-sound chart", however, does not exhaust the *hiragana* and *katakana* letters actually employed in Japanese, because the basic consonant sounds (*k, s, t, h*) have variants. The sound represented by the letter *h* is historically related to the sound represented by *p*, and these voiceless obstruents (*k, s, t*, and *p*) have their respective voiced counterparts (*g, z, d*, and *b*). Table 2 shows letters for these consonants followed by five vowels.

**Table 2:** *Letters for voiced obstruents and bilabial [p]*

| **transcription** | *ga* | *za* | *da* | *ba* | *pa* |
|---|---|---|---|---|---|
| *hiragana* | が | ざ | だ | ば | ぱ |
| *katakana* | ガ | ザ | ダ | バ | パ |
| **transcription** | *gi* | *zi* | *zi* | *bi* | *pi* |
| *hiragana* | ぎ | じ | ぢ | び | ぴ |
| *katakana* | ギ | ジ | ヂ | ビ | ピ |
| **transcription** | *gu* | *zu* | *zu* | *bu* | *pu* |
| *hiragana* | ぐ | ず | づ | ぶ | ぷ |
| *katakana* | グ | ズ | ヅ | ブ | プ |
| **transcription** | *ge* | *ze* | *de* | *be* | *pe* |
| *hiragana* | げ | ぜ | で | べ | ぺ |
| *katakana* | ゲ | ゼ | デ | ベ | ペ |
| **transcription** | *go* | *zo* | *do* | *bo* | *po* |
| *hiragana* | ご | ぞ | ど | ぼ | ぽ |
| *katakana* | ゴ | ゾ | ド | ボ | ポ |

It is important to note that Tables 1 and 2 show the conventional letters and alphabetical transcription adopted by the HJLL series; they are not intended to represent the actual pronunciations of Japanese vowels and consonants. For example, among the vowels, the sound represented as "u" is pronounced as [ɯ] with unrounded lips. Consonants may change articulation according to the following vowels. Romanization of these has been controversial with several competing proposals.

There are two Romanization systems widely used in Japan. One known as the Hepburn system is more widely used in public places throughout Japan such as train stations, street signs, as well as in some textbooks for learners of Japanese. This system is ostensibly easier for foreigners familiar with the English spelling system. The *Kunreishiki* (the cabinet ordinance system) is phonemic in nature and is used by many professional linguists. The essential differences between the two Romanization systems center on palatalized and affricate consonants, as shown in Table 3 below by some representative syllables for which two Romanization renditions differ:

**Table 3:** *Two systems of Romanization*

| Hiragana | IPA | Hepburn | Kunreishiki |
|---|---|---|---|
| し | [ʃi] | shi | si |
| しゃ | [ʃa] | sha | sya |
| しゅ | [ʃɯ] | shu | syu |
| しょ | [ʃo] | sho | syo |
| じ and ぢ | [ʤi] | ji | zi |
| じゃ | [ʤa] | ja | zya |
| じゅ | [ʤɯ] | ju | zyu |
| じょ | [ʤo] | jo | zyo |
| ち | [tʃi] | chi | ti |
| ちゃ | [tʃa] | cha | tya |
| ちゅ | [tʃɯ] | chu | tyu |
| ちょ | [tʃo] | cho | tyo |
| つ | [tsɯ] | tsu | tu |
| づ and ず | [dzɯ] | dzu | zu |
| ふ | [ɸɯ] | fu | hu |

Except for the volumes on Ryukyuan, Ainu, and Japanese dialects, whose phonetics differ from Standard Japanese, HJLL adopts the Kunreishiki system for rendering cited Japanese words and sentences but uses the Hepburn system for rendering conventional forms such as proper nouns and technical linguistic terms in the text and in the translations of examples.

The cited Japanese sentences in HJLL look as below, where the first line transliterates a Japanese sentence in Kunreishiki Romanization, the second line contains interlinear glosses largely following the Leipzig abbreviation convention, and the third line is a free translation of the example sentence.

(1) *Taroo wa Ziroo to Tookyoo e it-te kutusita o kat-ta.*
Taro TOP Jiro COM Tokyo ALL go-GER sock ACC buy-PST
'Taro went to Tokyo with Jiro and bought socks.'

The orthographic convention of rendering Japanese is to represent a sentence with an uninterrupted sequence of Sino-Japanese characters and *katakana* or *hiragana* syllabaries without a space for word segmentation, as in 太郎は次郎と東京へ行って靴下を買った for (1). In line with the general rules of Romanization adopted in

books and articles dealing with Japanese, however, HJLL transliterates example sentences by separating word units by spaces. The example in (1) thus has 10 words. Moreover, as in *it-te* (go-GERUNDIVE) and *kat-ta* (buy-PAST) in (1), word-internal morphemes are separated by a hyphen whenever necessary, although this practice is not adopted consistently in all of the HJLL volumes. Special attention should be paid to particles like *wa* (topic), *to* 'with' and *e* 'to, toward', which, in the HJLL representation, are separated from the preceding noun or noun phrase by a space (see section 7.3). Remember that case and other kinds of particles, though spaced, form phrasal units with their preceding nouns.

## 7.2 Word order

As seen in (1), Japanese is a verb-final, dependent-marking agglutinative language. It is basically an SOV language, which marks the nominal dependent arguments by particles (*wa*, *to*, *e*, and *o* above), and whose predicative component consists of a verbal-stem, a variety of suffixes, auxiliary verbs, and semi-independent predicate extenders pertaining to the speech act of predication (see section 7.6). While a verb is rigidly fixed in sentence final position, the order of subject and object arguments may vary depending on pragmatic factors such as emphasis, background information, and cohesion. Thus, sentence (2a) with the unmarked order below, in principle, may vary in multiple ways as shown by some possibilities in (2b)–(2d).

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (2) | a. | *Taroo* | *ga* | *Hanako* | *ni* | *Ziroo* | *o* | *syookai-si-ta.* |
| | | Taro | NOM | Hanako | DAT | Jiro | ACC | introducing-do-PST |
| | | 'Taro introduced Jiro to Hanako.' | | | | | | |
| | b. | *Taroo ga* ***Ziroo o*** *Hanako ni syookai-si-ta.* | | | | | | |
| | c. | ***Hanako ni*** *Taroo ga Ziroo o syookai-si-ta.* | | | | | | |
| | d. | ***Ziroo o*** *Taroo ga Hanako ni syookai-si-ta.* | | | | | | |

Adverbs, likewise, can be rather freely placed, though each type of adverbs has its basic position.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (3) | a. | ***Saiwainimo*** | *Hanako* | *ga* | *gohan* | *o* | *tai-te* | *kure-te* | *i-ta.* |
| | | luckily | Hanako | NOM | rice | ACC | cook-GER | GIVE-GER | BE-PST |
| | | 'Luckily Hanako had done the favor of cooking the rice (for us).' | | | | | | | |
| | b. | *Hanako ga* ***saiwainimo*** *gohan o tai-te kure-te i-ta.* | | | | | | | |
| | c. | *Hanako ga gohan o* ***saiwainimo*** *tai-te kure-te i-ta.* | | | | | | | |

Notice that while the verbal complex in the sentence above is not as tightly organized as a complex involving suffixes, a sentence adverb cannot be placed within the verbal complex, showing that the sequence of *tai-te kure-te i-ta* forms a tighter constituent,

which, however, permits insertion of the topic particle *wa* after each of the gerundive forms. (See section 7.4 below on the nature of gerundive forms in Japanese.)

As the normal position of sentence adverbs is sentence initial, manner and resultative adverbs have an iconically-motivated position, namely before and after the object noun phrase, respectively, as below, though again these adverbs may move around with varying degrees of naturalness:

(4) *Hanako ga* ***isoide*** *gohan o tai-te kure-ta.*
Hanako NOM hurriedly rice ACC cook-GER GIVE-PST
'Hanako did the favor of cooking the rice hurriedly (for us).'

(5) *Hanako ga gohan o* ***yawarakaku*** *tai-te kure-ta.*
Hanako NOM rice ACC softly cook-GER GIVE-PST
'Hanako did the favor of cooking the rice soft (for us).'

The fact that an object noun phrase can be easily separated from the verb, as in (2b.d), and that adverbs can freely intervene between an object and a verb, as in (5), has raised the question whether Japanese has a verb phrase consisting of a verb and an object noun phrase as a tightly integrated constituent parallel to the VP in English (cf. **cook hurriedly the rice* – the asterisk marks ungrammatical forms).

## 7.3 NP structure

Noun phrases, when they occur as arguments or adjuncts, are marked by case particles or postpositions that are placed after their host nouns. Because case markers can be set off by a pause, a filler, or even longer parenthetic material, it is clear that they are unlike declensional affixes in inflectional languages like German or Russian. Their exact status, however, is controversial; some researchers regard them as clitics and others as (non-independent) words.

Elaboration of Japanese noun phrases is done by prenominal modifiers such as a demonstrative, a genitive noun phrase, or an adjective, as below, indicating that Japanese is a consistent head-final language at both nominal and clausal levels.

(6) a. *kono Taroo no kaban*
this Taro GEN bag
lit. 'this Taro's bag'

b. *Taroo no kono kaban*
Taro GEN this bag
lit. 'Taro's this bag'

Japanese lacks determiners of the English type that "close off" NP expansion. The literal translations of the Japanese forms above are ungrammatical, indicating that English determiners like demonstratives and genitive noun phrases do not allow further expansion of an NP structure. Also seen above is the possibility that prenominal modifiers can be reordered just like the dependents at the sentence level. The order of prenominal modifiers, however, is regulated by the iconic principle of placing closer to the head noun those modifiers that have a greater contribution in specifying the nature and type of the referent. Thus, descriptive adjectives tend to be placed closer to a head noun than demonstratives and genitive modifiers of non-descriptive types. Interesting is the pattern of genitive modifiers, some of which are more descriptive and are placed closer to the head noun than others. Genitives of the same semantic type, on the other hand, can be freely reordered. Compare:

(7) a. *Yamada-sensei* *no* *kuroi* *kaban*
Yamada-professor GEN black bag
'Professor Yamada's black bag'
b. **kuroi Yamada-sensei no kaban*
(O.K. with the reading of 'a bag of Professor Yamada who is black')

(8) a. *Yamada-sensei* *no* *gengogaku* *no* *koogi*
Yamada-professor GEN linguistics GEN lecture
'Professor Yamada's linguistics lecture'
b. **gengogaku no Yamada-sensei no koogi*
(O.K. with the reading of 'a lecture by Professor Yamada of linguistics')

(9) a. *Yamada-sensei* *no* *kinoo* *no* *koogi*
Yamada-professor GEN yesterday GEN lecture
lit. 'Professor Yamada's yesterday's lecture' 'Yesterday's lecture by Professor Yamada'
b. *Kinoo no Yamada-sensei no koogi*

(10) a. *oomori* *no* *sio-azi* *no* *raamen*
big.serving GEN salt-tasting GEN *ramen*
lit. 'big-serving salt-tasting *ramen* noodles'
b. *sio-azi no oomori no raamen*

(11) a. *atui* *sio-azi* *no* *raamen*
hot salt-tasting GEN *ramen*
'hot salt-tasting *ramen* noodles'
b. *sio-azi no atui ramen*

Numeral classifiers (CLFs) pattern together with descriptive modifiers so that they tend to occur closer to a head noun than a possessive genitive phrase.

(12) a.

| | | | | |
|---|---|---|---|---|
| *Taroo* | *no* | *san-bon* | *no* | *enpitu* |
| Taro | GEN | three-CLF | GEN | pencil |

‘Taro’s three pencils’

b. **san-bon no Taroo no enpitu*

Numeral classifiers also head an NP, where they play a referential function and where they can be modified by a genitive phrase or an appositive modifier, as in (13a.b). They may also “float” away from the head noun and become adverbial, as in (13c).

(13) a.

| | | | | | | |
|---|---|---|---|---|---|---|
| *Taroo* | *wa* | *gakusei* | *no* | ***san-nin*** | *o* | *mikake-ta.* |
| Taro | TOP | student | GEN | three-CLF | ACC | see.by.chance-PST |

‘Taro saw three of students by chance.’

b.

| | | | | | |
|---|---|---|---|---|---|
| *Taroo* | *wa* | *gakusei* | ***san-nin*** | *o* | *mikake-ta.* |
| Taro | TOP | student | three-CLF | ACC | see.by.chance-PST |

lit. ‘Taro saw student-threes by chance.’

c.

| | | | | | |
|---|---|---|---|---|---|
| *Taroo* | *wa* | *gakusei* | *o* | ***san-nin*** | *mikake-ta.* |
| Taro | TOP | student | ACC | three-CLF | see.by.chance-PST |

‘Taro saw students, three (of them), by chance.’

As in many other SOV languages, the so-called relative clauses are also prenominal and are directly placed before their head nouns without the mediation of “relative pronouns” like the English *which* or *who* or “complementizers” like *that*. The predicates in relative clauses are finite, taking a variety of tense and aspect. The subject may be replaced by a genitive modifier. Observe (14a).

(14) a.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Boku* | *mo* | *[Taroo* | *ga/no* | *kat-ta]* | *hon* | *o* | *kat-ta.* |
| I | ADVPART | Taro | NOM/GEN | buy-PST | book | ACC | buy-PST |

‘I also bought the book which Taro bought.’

b.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Boku* | *mo* | *[Taroo* | *ga/no* | *kat-ta]* | *no* | *o* | *kat-ta.* |
| I | ADVPART | Taro | NOM/GEN | buy-PST | NM | ACC | buy-PST |

‘I also bought the one which Taro bought.’

The structure used as a modifier in the relative clause construction can also head a noun phrase, where it has a referential function denoting an entity concept evoked by the structure. In Standard Japanese such a structure is marked by the nominalization particle *no*, as in (14b).

## 7.4 Subject and topic

Some of the sentences above have noun phrases marked by the nominative case particle *ga* and some by the topic marker *wa* for what appear to correspond to the subject noun phrases in the English translations. This possibility of *ga*- and *wa*-marking is seen below.

(15) a. *Yuki ga siro-i.*
snow NOM white-PRS
'The snow is white.'

b. *Yuki wa siro-i.*
snow TOP white-PRS
'Snow is white.'

As the difference in the English translations indicates, these two sentences are different in meaning. Describing the differences between topic and non-topic sentences has been a major challenge for Japanese grammarians and teachers of Japanese alike. The difference in the English translations above, however, is indicative of how these two sentences might differ in meaning. Sentence (15a) describes a state of affairs involving specific snow just witnessed, whereas (15b) is a generic statement about a property of snow unbounded by time. Thus, while (15a) would be uttered only when the witnessed snow is indeed white, (15b) would be construed true even though we know that there are snow piles that are quite dirty.

A similar difference is seen in verbal sentences as well.

(16) a. *Tori ga tob-u.*
bird NOM fly-PRS
'A bird is flying/is about to fly.'

b. *Tori wa tob-u.*
bird TOP fly-PRS
'Birds fly.'

Non-topic sentences like (15a) and (16a) are often uttered with an exclamation accompanying a sudden discovery of a state of affairs unfolding right in front of one's eyes. The present tense forms (*-i* for adjectives and *-(r)u* for verbs) here anchor the time of this discovery to the speech time. The present tense forms in (15b) and (16b), on the other hand, mark a generic tense associated with a universal statement.

These explanations can perhaps be extended to a time-bound topic sentence seen in (17b) below.

(17) a.

| | | |
|---|---|---|
| *Taroo* | *ga* | *hasit-ta.* |
| Taro | NOM | run-PST |

'Taro NOM ran.'

b.

| | | |
|---|---|---|
| *Taroo* | *wa* | *hasit-ta.* |
| Taro | TOP | run-PST |

'Taro ran.'

That is, while (17a) reports an occurrence of a particular event at a time prior to the speech time, (17b) describes the nature of the topic referent – that Taro was engaged in the running activity – as a universal truth of the referent, but universal only with respect to a specifically bound time marked by the past tense suffix.

Topics need not be a subject, and indeed any major sentence constituent, including adverbs, may be marked topic in Japanese, as shown below.

(18) a.

| | | | | | | |
|---|---|---|---|---|---|---|
| *Sono* | *hon* | *wa* | *Taroo* | *ga* | *yon-de* | *i-ru.* |
| that | book | TOP | Taro | NOM | read-GER | BE-PRS |

'As for that book, Taro is reading (it).'

b.

| | | | | |
|---|---|---|---|---|
| *Kyoo* | *wa* | *tenki* | *ga* | *yo-i.* |
| today | TOP | weather | NOM | good-PRS |

'As for today, the weather is good.'

c.

| | | | | | |
|---|---|---|---|---|---|
| *Sonnani* | *wa* | *hayaku* | *wa* | *hasir-e* | *na-i.* |
| that.way | TOP | quickly | TOP | run-POTEN | NEG-PRS |

'That quickly, (I) cannot run.'

## 7.4 Complex sentences

As in many Altaic languages, compound sentences in Japanese do not involve a coordinate conjunction like English *and*. Instead, clauses are connected by the use of inflected verb forms, as in (19a) below, where the *-i* ending is glossed in the HJLL series as either INF (infinitive) or ADVL (adverbal) following the Japanese term *ren'yō-kei* for the form. While the *-i* ending in the formation of compound sentences is still used today, especially in writing, the more commonly used contemporary form involves a conjunctive particle *-te* following the *-i* infinitive form, as in (19b) below. In HJLL, this combination is glossed as GER (gerundive), though the relevant Japanese forms do not have the major nominal use of English gerundive forms.

(19) a.

| | | | | | |
|---|---|---|---|---|---|
| *Hana* | *wa* | *sak-i,* | *tori* | *wa* | *uta-u.* |
| flower | TOP | bloom-INF | bird | TOP | sing-PRS |

'Flowers bloom and birds sing.'

b. *Hana wa sa.i-te, tori wa uta-u.*
flower TOP bloom-GER bird TOP sing-PRS
'Flowers bloom and birds sing.'

Both the *-i* and *-te* forms play important roles in Japanese grammar. They are also used in clause-chaining constructions for serial events (20a), and in complex sentences (20b)–(20d), as well as in numerous compound verbs (and also in many compound nouns) such as *sak-i hokoru* (bloom-INF boast) 'be in full bloom', *sak-i tuzukeru* (bloom-INF continue) 'continue blooming', *sa.i-te iru* (bloom-GER BE) 'is blooming', and *sa.i-te kureru* (bloom-GER GIVE) 'do the favor of blooming (for me/us)'.

(20) a. *Taroo wa [ok-i/ok.i-te], [kao o ara-i/arat-te],*
Taro TOP rise-INF/rise-GER face ACC wash-INF/wash-GER
*[gohan o tabe-ta].*
meal ACC eat.PST
'Taro got up, washed his face, and ate a meal.'

b. *Taroo wa [sakana o tur-i] ni it-ta.*
Taro TOP fish ACC catch-INF DAT go-PST
'Taro went to catch fish.'

c. *Taroo wa [aruk-i nagara] hon o yon-da.*
Taro TOP walk-INF SIMUL book ACC read-PST
'Taro read a book while walking.'

d. *Taroo wa [Hanako ga ki-ta no] ni awa-na-katta.*
Taro TOP Hanako NOM come-PST NM DAT see-NEG-PST.
'Taro did not see (her), even though Hanako came.'

(20d) has the nominalized clause marked by the particle *no* followed by the dative *ni*, also seen in (20b) marking the purposive form. Now the *no-ni* sequence has been reanalyzed as a concessive conjunction meaning 'even though'.

## 7.5 Context dependency

The context dependency of sentence structure in Japanese is much more clearly pronounced than in languages like English. Indeed, it is rare that Japanese sentences express all the arguments of a verb such as a subject (or topic) and an object noun phrase included in the sentences used above for illustrative purposes. A typical dialog would take the following form, where what is inferable from the speech context is not expressed.

(21) a. Speaker A: *Tokorode, Murakami Haruki no saisin-saku yon-da ka.*
by.the.way Murakami Haruki GEN newest-work read-PST Q
'By the way, have (you) read Haruki Murakami's latest work?'

b. Speaker B: *Un, moo yon-da.*
uh-hu already read-PST
'Uh-hu, (I) already read (it)'.

In (21a) A's utterance is missing a subject noun phrase referring to the addressee, and B's response in (21b) is missing both subject and object noun phrases. In some frameworks, sentences like these are analyzed as containing zero pronouns or as involving a process of "pro drop", which deletes assumed underlying pronouns. This kind of analysis, however, ignores the role of speech context completely and incorporates information contextually available into sentence structure. In an analysis that takes seriously the dialogic relationship between speech context and sentence structure, the expressions in (21) would be considered full sentences as they are.

## 7.6 Predicative verbal complexes and extenders

Coding or repeating contextually determinable verb phrases, as in (21b), is less offensive than expressing contextually inferable noun phrases presumably because verb phrases have the predication function of assertion, and because they also code a wide range of other types of speech acts and of contextual information pertaining to the predication act. Declarative sentences with plain verbal endings like the one in (21b) are usable as "neutral" expressions in newspaper articles and literary works, where no specific reader is intended. In daily discourse, the plain verbal forms "explicitly" code the speaker's attitude toward the hearer; namely, that the speaker is treating the hearer as his equal or inferior in social standing, determined primarily by age, power, and familiarity. If the addressee were socially superior or if the occasion demanded formality, a polite, addressee honorific form with the suffix *-masu* would be used, as below.

(22) *Hai, moo yom-i-masi-ta.*
yes already read-INF-POL-PST
'Yes, (I have) already read (it).'

The referent honorific forms are used when the speaker wishes to show deference toward the referent of arguments – subject honorific and object honorific (or humbling) forms depending on the type of argument targeted. If (21b) were to be uttered in reference to a social superior, the following would be more appropriate:

(23) *Un, (Yamada-sensei wa) moo yom-are-ta.*
uh-hu (Yamada-professor TOP) already read-SUB.HON-PST
'Uh-hu, (Professor Yamada has) already read (it).'

This can be combined with the polite ending *-masu*, as below, where the speaker's deference is shown to both the referent of the subject noun phrase and the addressee:

(24) *Hai, (Yamada-sensei wa) moo yom-are-masi-ta.*
Yes (Yamada-professor TOP) already read-HON-POL-PST
'Yes, (Professor Yamada has) already read (it).'

As these examples show, Japanese typically employs agglutinative suffixes in the elaboration of verbal meanings associated with a predication act. The equivalents of English auxiliary verbs are either suffixes or formatives connected to verb stems and suffixed forms in varying degrees of tightness. These are hierarchically structured in a manner that expresses progressively more subjective and interpersonal meaning as one moves away from the verb-stem core toward the periphery. For example, in the following sentence a hyphen marks suffixal elements tightly bonded to the preceding form, an equal sign marks a more loosely connected formative, which permits insertion of certain elements such as the topic particle *wa*, and a space sets off those elements that are independent words following a finite predicate form, which may terminate the utterance.

(25) *(Taroo wa) ik-ase-rare-taku=na-katta rasi-i mitai des-u wa.*
(Taro TOP) go-CAUS-PASS-DESI=NEG-PST CONJEC-PRS UNCERT POLCOP-PRS SFP
'(Taro) appears to seem to not want to have been forced to go, I tell you.'

The final particle *wa* above encodes the information that the speaker is female. A male speaker would use *yo* or *da yo*, the latter a combination of the plain copula and *yo*, instead of *desu wa* above, or combinations such as *da ze* and *da zo* in rough speech.

Non-declarative Japanese sentences, on the other hand, frequently suppress auxiliary verbs, the copula, and the question particle especially in casual speech, where intonation and tone of voice provide clues in guessing the intended speech act. Casual interrogatives take the form of (26a) with a nominalization marker bearing a rising intonation, marked by the question mark in the transcription, whereas fuller versions have the interrogative particle *ka* or a combination of the polite copula and *ka*, as in (26b).

(26) a. *Moo kaeru no?*
already return NM
'Going home already?'

b. *Moo kaeru no (desu) ka.*
already return NM (POLCOP) Q
'Going home already?'

Requests are made with the aid of an auxiliary-like "supporting" verb *kureru* 'GIVE (ME THE FAVOR OF. . .)', its polite form *kudasai*, or its intimate version *tyoodai*, as seen in (27a). Again, these forms are often suppressed in a highly intimate conversation and may result in a form like (27b).

(27) a. *Hayaku kaet-te kure/kudasai/tyoodai.*
soon return-GER GIVE/GIVE.POL/GIVE.INTI
'(Please) come home soon (for me/us).'

b. *Hayaku kaet-te ne.*
soon return-GER SFP
'(Please) come home soon, won't you?'

The use of dependent forms (e.g., the gerundive *-te* form above) as independent sentences is similar to that of subjunctive forms of European languages as independent sentences, as illustrated by the English sentence below.

(28) *If you would give me five thirty-cent stamps.*

Conditionals are used as independent suggestion sentences in Japanese as well. For example, (29a) has a fuller version like (29b) with the copula as a main-clause verb, which can also be suppressed giving rise to the truncated form (29c).

(29) a. *Hayaku kaet-tara?*
quickly return-COND
lit. ' If return quickly.' 'Why don't you go home quickly?'

b. *Hayaku kaet-tara ikaga desu ka.*
quickly return-COND how POLCOP Q
lit. 'How is it if (you) went home quickly?'

c. *Hayaku kaet-tara ikaga?*
quickly return-COND how
'Why don't (you) go home quickly?'

Understanding Japanese utterances requires full recourse to the elements of speech context, such as the nature of the speaker and the hearer and the social relationship between them, the information "in the air" that is readily accessible to the interlocutors, and the formality of the occasion. Indeed, the difficult part of the art of

speaking Japanese is knowing how much to leave out from the utterance and how to infer what is left unsaid.

## 8 Conclusion

Many of the interesting topics in Japanese grammar introduced above are discussed in great detail in the Lexicon-Word formation handbook and the Syntax volume. The Historical handbook also traces developments of some of the forms and constructions introduced above. The Sociolinguistics volume gives fuller accounts of the sentence variations motivated by context and discourse genre.

## References

Chamberlain, Basil H. 1895. *Essay in aid of a grammar and dictionary of the Luchuan language*. Transactions of the Asiatic Society of Japanese, vol. 23 supplement.
Frellesvig, Bjarke. 2010. *A history of the Japanese language*. Cambridge: Cambridge University Press.
Martin, Samuel E. 1975. *A reference grammar of Japanese*. New Haven: Yale University Press.
Miller, Roy A. 1971. *Japanese and the other Altaic languages*. Chicago: University of Chicago Press.
Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.

## Appendix: List of abbreviations for HJLL

| | |
|---|---|
| 1 | first person |
| 2 | second person |
| 3 | third person |
| A | agent-like argument of canonical transitive verb |
| ABL | ablative |
| ACC | accusative |
| ACOP | adjectival copula |
| ADJ | adjective |
| AND | adnominal |
| ADV | adverb(ial(izer)) |
| ADVL | adverbal |
| ADVPART | adverbial particle |
| AGR | agreement |
| AGT | agent |
| ALL | allative |
| AN | adjectival noun |

| | |
|---|---|
| ANTIP | antipassive |
| AP | adverbial particle, adjective phrase |
| APPL | applicative |
| ART | article |
| ASP | aspect |
| ATTR | attributive |
| AUX | auxiliary |
| AUXV | auxiliary verb |
| C | consonant |
| CAUS | causative |
| CLF | classifier |
| COHORT | cohortative |
| COM | comitative |
| COMP | complementizer |
| COMPL | completive |
| CONC | concessive |
| CONCL | conclusive |
| COND | conditional |
| CONJEC | conjectural |
| CONJCT | conjunctive |
| CONT | continuative |
| COP | copula |
| CVB | converb |
| DAT | dative |
| D | demonstrative |
| DECL | declarative |
| DEF | definite |
| DEM | demonstrative |
| DET | determiner |
| DESI | desiderative |
| DIST | distal |
| DISTR | distributive |
| DO | direct object |
| DU | dual |
| DUR | durative |
| EMPH | emphatic |
| ERG | ergative |
| ETOP | emphatic topic |
| EVID | evidential |
| EXCL | exclamatory, exclusive |
| EXPL | expletive |
| FOC | focus |

| | |
|---|---|
| FUT | future |
| GEN | genitive |
| GER | gerund(ive) |
| H | high (tone or pitch) |
| HON | honorific |
| HUM | humble |
| IMP | imperative |
| INCL | inclusive |
| IND | indicative |
| INDEF | indefinite |
| INF | infinitive |
| INS | instrumental |
| INT | intentional |
| INTERJEC | interjection |
| INTI | intimate |
| INTR | intransitive |
| IO | indirect object |
| IRR | irrealis |
| ITERA | iterative |
| k-irr | k-irregular (*ka-hen*) |
| L | low (tone or pitch) |
| LB | lower bigrade (*shimo nidan*) |
| LM | lower monograde (*shimo ichidan*) |
| LOC | locative |
| MPST | modal past |
| MVR | mid vowel raising |
| N | noun |
| n-irr | n-irregular (*na-hen*) |
| NCONJ | negative conjectual |
| NEC | neccessitive |
| NEG | negative |
| NM | nominalization marker |
| NMLZ | nominalization/nominalizer |
| NMNL | nominal |
| NOM | nominative |
| NONPST | nonpast |
| NP | noun phrase |
| OBJ | object |
| OBL | oblique |
| OPT | optative |
| P | patient-like argument of canonical transitive verb, preposition, post-position |

| | |
|---|---|
| PART | particle |
| PASS | passive |
| PCONJ | present conjectural |
| PERF | perfective |
| PL | plural |
| POL | polite |
| POLCOP | polite copula |
| POSS | possessive |
| POTEN | potential |
| PP | prepositional/postpositional phrase |
| PRED | predicative |
| PRF | perfect |
| PRS | present |
| PRES | presumptive |
| PROG | progressive |
| PROH | prohibitive |
| PROV | provisional |
| PROX | proximal/proximate |
| PST | past |
| PSTCONJ | past conjectural |
| PTCP | participle |
| PURP | purposive |
| Q | question/question particle/question marker |
| QD | quadrigrade (*yodan*) |
| QUOT | quotative |
| r-irr | r-irregular (*ra-hen*) |
| REAL | realis |
| RECP | reciprocal |
| REFL | reflexive |
| RES | resultative |
| RESP | respect |
| S | single argument of canonical intransitive verb, sentence |
| SBJ | subject |
| SBJV | subjunctive |
| SFP | sentence final particle |
| SG | singular |
| SIMUL | simultaneous |
| s-irr | s-irregular (*sa-hen*) |
| SG | singular |
| SPON | spontaneous |
| SPST | simple past |
| STAT | stative |

| | |
|---|---|
| TOP | topic |
| TR | transitive |
| UB | upper bigrade (*kami-nidan*) |
| UNCERT | uncertain |
| UM | upper monograde (*kami-ichidan*) |
| V | verb, vowel |
| VN | verbal noun |
| VOC | vocative |
| VOL | volitional |
| VP | verb phrase |

## Languages

| | |
|---|---|
| ConJ | contemporary Japanese |
| EMC | Early Middle Chinese |
| EMJ | Early Middle Japanese |
| EOJ | Eastern Old Japanese |
| J-Ch | Japano-Chinese |
| LMC | Late Middle Chinese |
| LMJ | Late Middle Japanese |
| JPN | Japanese |
| MC | Middle Chinese |
| MJ | Middle Japanese |
| MK | Middle Korean |
| ModJ | Modern Japanese |
| OC | Old Chinese |
| OJ | Old Japanese |
| pJ | proto-Japanese |
| pK | proto-Korean |
| SJ | Sino-Japanese |
| Skt | Sanskrit |

# In Memory of Tsutomu Sakamoto [1954–2014]

While working on this handbook, we received the sad news that Tsutomu Sakamoto had passed away in Fukuoka, Japan.

Tsutomu was a theoretical linguist, and a trained psycholinguist, specializing in metaphor and sentence processing. After obtaining his Ph. D. from the City University of New York, he taught at Kobe Shoin Women's University and Kyushu University. He was an active member/contributor of the Japanese Cognitive Science Society and the Linguistics Society of Japan. He was loved by his colleagues and students at the institutions where he taught. I have known him almost 30 years. He was a caring, open-minded and down-to-earth kind of person. I personally have many fond memories of him: he hosted me in his New York City apartment, he took me fishing in Genkainada, he introduced me to his colleagues to start a study abroad program, he showed me his lab and most of all we discussed many linguistic issues with his fellow graduate students. He was a great friend and an excellent colleague. Many of us have personal stories with him and will treasure them. We all miss him very much. This unfortunate incident happened when he was expected to contribute even more to the field. His work on sentence processing tackles many interesting issues in both theoretical and experimental linguistics, so our loss is great. Certainly his work will continue to influence the students of Japanese linguistics for years to come. May Tsutomu rest in peace in Genkainada.

This handbook is dedicated to the late psycholinguist, Tsutomu Sakamoto. We are very fortunate to be able to include his last article, "Processing of syntactic and semantic information in the human brain: Evidence from ERP studies in Japanese", as Chapter 15.

Mineharu Nakayama
Handbook of Japanese Psycholinguistics, Editor
Columbus, Ohio, U.S.A.

# Acknowledgments

First, I would like to thank the General Series Editors, Masayoshi Shibatani and Taro Kageyama, for the opportunity they provided me with this project and their valuable assistance during the editorial process including comments on the volume chapters. I also extend my gratitude to English and style check editors John Haig and Yukie Sato for their careful assistance.

Of course, this handbook could not have reached the final form without various people's professional contributions. I would like to thank the contributors of the chapters and the reviewers who have offered critical and constructive comments that greatly enhanced the quality of the final product. Some of them read multiple chapters. By listing their names, I thank them for their extensive and thoughtful comments for this publication: Brian Baker, Charles Bowman, Erik Christensen, Hannah Dodd, Rachel Fatica, Seth Goss, John Grinstead, Akiko Kashiwagi-Wood, Hsing-Yin Lin, Laura Maurer, Nan Meng, Masahiko Minami, Jennifer Nakayama, Koichi Sawasaki, Keiko Schaadt, Tomohiko Shirahata, Yi-Ching Su, Natsumi Suzuki, Shinsuke Tsuchiya, Laura Wagner, Zhiguo Xie, and Noriko Yoshimura. I would also like to thank Emily Farrell, Lara Wysong, Wolfgang Konwitschy and Uri Tadmor of De Gruyter Mouton for their assistance.

Finally, I thank the Department of Linguistic Theory and Structure (Director: Haruo Kubozono) at the National Institute for Japanese Language and Linguistics, and the Department of East Asian Languages and Literatures, the Institute for Japanese Studies, and the College of Arts and Sciences at The Ohio State University for their generous assistance. Without it, this publication would not have been possible and I hereby acknowledge the support of the National Institute for Japanese Language and Linguistics and The Ohio State University throughout the project.

Mineharu Nakayama

# Table of contents

# Contributors

**Franklin Chang** is a lecturer in the Department of Psychology, University of Liverpool, UK. His research examines the relationship between production and learning through connectionist models and human experiments. He is interested in cross-linguistic differences in acquisition or processing and has done studies in English, German, Korean, and Japanese.

**Shinji Fukuda** is an associate professor in the Faculty of Psychological Science, Health Sciences University of Hokkaido, Japan. His primary research interests are the acquisition of syntactic and morphological properties in Japanese children with both specific language impairment and normal language development, and their neurological implications.

**Suzy E. Fukuda** is a professor of English in the Faculty of Law, Aoyama Gakuin University, Japan. Her primary research interests are the acquisition of phonology and morphology in Japanese children with both specific language impairment and normal language development, and their underlying neurological properties.

**Allison Gabriele** is an associate professor in the Department of Linguistics, University of Kansas, USA. Her primary research interest is in second language acquisition with a focus on syntax and semantics. Her research uses both psycholinguistic and neurolinguistic techniques to examine the role of both linguistic and cognitive factors in second language development and processing. She is an Associate Editor for the journal *Linguistic Approaches to Bilingualism*.

**Takuya Goro** is an associate professor in the Department of English, Tsuda College, Japan. His research interest lies in first language acquisition, especially, on quantification and logical expressions.

**Hiroko Hagiwara** is a professor in the Department of Language Sciences and the director of the Research Center for Brain, Language and Genetics at Tokyo Metropolitan University, Japan. She has been an editorial board member of various international journals. Her research covers a wide range of topics in neurolinguisitics and cognitive neuroscience of language, including first and second/foreign language acquisition/learning and processing in Japanese normal populations and language breakdown of Japanese aphasic patients. Her current research explores the neural basis of the acquisition of syntax and semantics in children and adolescents.

**Yuki Hirose** is an associate professor in the Department of Language and Information Sciences, University of Tokyo, Japan. Her research interests are in human sentence processing. Her recent research investigates how human listeners interpret prosodic information on line and how that process interacts with various other processing decisions (syntactic structure, word boundary, referential status, etc.) in Japanese.

**Mutsumi Imai** is a professor of psychology at the Faculty of Environment and Information Studies, Keio University, Japan. Her research interests include three questions that are deeply interwoven with one another, how children learn language, the role of language in conceptual development, and the relationship between language and thought (in cross linguistic contexts) in children and adults. She has received awards from the Japanese Cognitive Science Society, the Japanese Society of Human Development, and the Japanese Psychological Society for her distinguished academic achievements and international contributions. She has been an associate member of the Science Council of Japan since 2007, and served on government committees concerning scientific research and education.

**Tomohiko Ito** is a professor in the Faculty of Education for Children with Disabilities, Tokyo Gakugei University, Japan. His primary research interests are the acquisition and disorders of language in Japanese children, especially the relationship between the acquisition of syntax and the onset of stuttering in young children.

**Noriko Iwasaki** is a senior lecturer in the Department of Linguistics, SOAS (School of Oriental and African Studies), University of London, UK. Her research interest lies in L2 production, (particularly grammatical encoding), L2 acquisition and language pedagogy.

**Barış Kahraman** is a post-doc research fellow of the Japan Society for the Promotion of Science at the University of Tokyo, Japan. His main research interests are human sentence processing, second language acquisition, teaching Japanese as a second language and Japanese-Turkish linguistics.

**Junko Kanero** is a doctoral student in the Department of Psychology, Temple University, USA. Her research interests include language development in infancy and childhood, neural processing of language, and the relationship between language and non-linguistic cognition. With a broad range of methodologies (behavioral measures, eye tracking, EEG, and fMRI), her research examines language as a window into human cognition.

**Akiko Kashiwagi-Wood** is an assistant professor in the Department of Modern Languages and Literatures, Oakland University, USA. Her research interests are sentence processing in a foreign language and second language acquisition. She is particularly interested in whether non-native speakers process sentences differently from native speakers and whether this may influence the ultimate attainment of the second language.

**Masatoshi Koizumi** is an associate professor in the Department of Linguistics, Tohoku University, Japan. His research interests are in grammatical theory and neurocognition of language. He is the author of *Phrase Structure in Minimalist Syntax* (Hituzi Syobo, 1999) and a coauthor of *Bun-no Kōzō* [Clausal Architecture] (with Koichi Tateishi, Kenkyusha, 2001).

**Reiko Mazuka** is a laboratory head at the Laboratory for Language Development at RIKEN Brain Science Institute in Japan, and a research professor in the Department of Psychology and Neuroscience at Duke University, USA. Her earlier research focused on Japanese sentence processing by children and adults. For the past 10 years, her research has focused on the acquisition of phonology by Japanese infants.

**Masahiko Minami** is a professor at San Francisco State University, USA, where he specializes in Japanese language and cross-cultural studies. He also serves as an invited professor at the National Institute for Japanese Language and Linguistics, Tokyo, Japan. He is currently Editor-in-Chief of the *Journal of Japanese Linguistics*; he also edits the *Handbook of Japanese Applied Linguistics* in the Handbooks of Japanese Language and Linguistics series. His recent works include *Telling Stories in Two Languages: Multiple Approaches to Understanding English-Japanese Bilingual Children's Narratives* (Information Age Publishing, 2011).

**Keiko Murasugi** is a professor of Linguistics at the Faculty of Foreign Studies/ Center for Linguistics, Nanzan University, Japan. Her main research interests are the analysis of adult syntax and first-language acquisition of Japanese (and English) from the perspective of Generative Grammar. Her research focuses on the emerging morphosyntax, and children's "errors" that occur within the limits of Universal Grammar.

**Mineharu Nakayama** is a professor in the Department of East Asian Languages and Literatures at The Ohio State University, USA, as well as a visiting professor at the National Institute for Japanese Language and Linguistics, Tokyo, Japan. His research interests are in L1 and L2 language acquisition and language processing. He is a former editor-in-chief of the *Journal of Japanese Linguistics*, an author of *Acquisition of Japanese Empty Categories* (Kurosio Publishers, 1996) and a co-editor of the *Handbook of East Asian Psycholinguistics, Vol. 2, Japanese* (with Reiko Mazuka and Yasuhiro Shirai, 2006).

**Hiromu Sakai** is a professor in the Graduate School of Education and the director of the Center for Brain Science of Language and Cognition, Hiroshima University, Japan. His main research interests are theoretical linguistics, psycholinguistics, and cognitive neuroscience of language.

**Tsutomu Sakamoto** was a professor in the Department of Linguistics, Kyushu University, Japan. His research interest was psycholinguistics, especially, sentence comprehension. He authored *Processing Empty Subjects in Japanese: Implications for the Transparency Hypothesis* (Kyushu University Press, 1996), and co-authored *Gengokagaku to kanren ryouiki* [Linguistic Science and its Related Research Area] (with Yukio Otsu, Toshiro Inui, Yoshihiro Nishimitu, and Nobuo Okada, Iwanami Shoten, 1998), and *Gengogaku no hōhō* [Method in Linguistics: Introduction to Modern Linguistics vol. 1] (with Takao Gunji, Iwanami Shoten, 1999). He also edited *Communicating Skills of Intention*. (Hituzi Syobo, 2007).

**Koichi Sawasaki** is an associate professor in the School of International Relations, University of Shizuoka, Japan. His research interests are Japanese sentence processing and second language acquisition. His recent articles include “The role of reading experience in construing Japanese gapless relative clauses” (2014), “Dokusho keiken to sadō kioku ga bunrikai ni oyobosu eikyō [Reading experience and working memory, and their effects on sentence comprehension]” (2012), and “Processing of relative clauses by learners of Japanese: A study on reading times of English/Korean/Chinese L1 speakers” (2009).

**Yasuhiro Shirai** is a professor of Applied Linguistics and Eirik Borve Chair in Modern Languages at the Department of Modern Languages and Literatures in Case Western Reserve University. His research interests lie in L1 and L2 acquisition, in particular, the cross-linguistic study of tense-aspect morphology. He authored *Second Language Acquisition: An Introduction for English Language Teachers* (Taishukan, 2012) among others, and co-authored *The Acquisition of Lexical and Grammatical Aspect* (with P. Li, Mouton de Gruyter, 2000).

**Mamori Sugita Hughes** is an adjunct assistant professor of Japanese at Queens College, City University of New York, USA. Her primary research interests are in the syntax/semantics interface, focusing on the syntax and semantics of aspect in Japanese. She also has a secondary interest in second language acquisition and foreign language teaching. She has written several self-study books published by Living Language/Random House for the study of Japanese.

**Katsuo Tamaoka** is a professor at the Graduate School of Languages and Cultures, Nagoya University, Japan. His research covers a wide range of issues in language processing – from orthography and words to phrasal and sentence-level topics. He is the editor of a special issue of *Japanese Psychological Research* (*The production of speech sounds across languages*). His articles appear in journals such as the *Journal of Neurolinguistics, Journal of Psycholinguistic Research, Psychonomic Bulletin and Review*, the *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *Language*, *Linguistic Inquiry*, and *Journal of Japanese Linguistics*.

**Noriko Yoshimura** is a professor in the School of International Relations, University of Shizuoka, Japan. Her teaching and research focuses on the relationship between linguistic theory and language acquisition. She co-authored *Kaigai tanki eigo kenshū to dai 2 gengo shūtoku* [Short-term English study abroad and L2 acquisition] (with Mineharu Nakayama, Hituzi Syobo, 2010).

## List of abbreviations

| | |
|---|---|
| ABS | absolutive |
| ACC | accusative |
| AdjP | adjectival phrase |
| ASP | aspect |
| AspP | aspect phrase |
| CAUS | causative |
| CLF | classifier |
| COMP | complementizer |
| COP | copula |
| CP | complementizer phrase |
| DAT | dative |
| DESD | desiderative |
| DET | determiner |
| DP | determiner phrase |
| ERG | ergative |
| GEN | genitive |
| GER | gerund(ive) |
| HON | honorific |
| IND | indicative |
| INF | infinitive |
| INTR | intransitive |
| IP | inflectional phrase/ intonation phrase |
| LOC | locative |
| MIM | mimetic word |
| MOD | modal |
| MP | modal phrase/major phrase |
| MOOD | mood marker |
| N | noun |
| NEG | negative |
| NMLZ | nominalization/nominalizer |
| NOM | nominative |
| NONPST | nonpast |
| NP | noun phrase |
| OBL | oblique |
| PASS | passive |
| PST | past |
| PERF | perfective |
| PL | plural |
| POL | polite |
| POTEN | potential |
| PP | prepositional/ postpositional phrase |
| PRS | present |
| PROG | progressive |
| PST | past |
| Q | question/question particle/ question marker |
| RES | resultative |
| REQ | request |
| SFP | sentence final particle |
| SG | singular |
| TOP | topic |
| TP | tense phrase |
| TR | transitive |
| V | verb, vowel |
| VP | verb phrase |

Mineharu Nakayama

# Japanese psycholinguistics and this volume

## 1 Introduction

Japanese is one of the most well-studied non-Indo-European languages in the field of linguistics. Two handbooks of Japanese linguistics have already been published in English (Tsujimura 1999; Miyagawa and Saito 2008), and this book is the second handbook to be published in English fully devoted to topics in Japanese psycholinguistics (the first book: Nakayama, Mazuka and Shirai 2006). Unlike the previous handbooks, however, the current handbook is slightly different, as it is one of the volumes in the Handbooks of Japanese Language and Linguistics (HJLL) series by the National Institute for Japanese Language and Linguistics and De Gruyter Mouton. Their differences will be described below.

This introductory chapter provides a general background of Japanese psycholinguistics, specifically, language acquisition and language processing. The next two sections discuss the backgrounds of general and Japanese psycholinguistics fields, outlining each of Japanese language acquisition and processing subfields. Then, the editorial background for this volume will be explained referring to the differences between the two Japanese psycholinguistics handbooks. The editor's final comments appear after describing each chapter of first language acquisition, second language acquisition, L1 Japanese processing and L2 Japanese processing.

## 2 Psycholinguistics

*Shinrigengogaku* is a Japanese term equivalent to *psycholinguistics*. However, the term *psycholinguistics* is sometimes translated into another term *gengoshinrigaku*. As both terms contain *shinri* 'psychology' and *gengo* 'language', scholars from two fields – psychology and linguistics – work in these fields although their approaches/focuses are possibly different. In Japanese, the word that appears just before –*gaku* 'study' determines the main focus or approach of the study. Herman (2004) calls the former *Psycholinguistik* and the latter *Sprachpsychologie* in German and differentiates them technically. The former studies a language as a universal system whereas the latter focuses on psychological and neurological functions in linguistic activities. Although the terminological differences come from their original definitions and history, their boundaries are, in reality, blurry and not clearly separated these days. People broadly use the term *shinrigengogaku* to cover both approaches. Sometimes even this term is very inclusive, including applied linguistics (educational linguistics), biolinguistics, computational linguistics, neurolinguistics, and so on.

Psycholinguistics is inherently an interdisciplinary field. As such, the theoretical approaches individual psycholinguists take in the field vary, though mostly they are from either linguistics or psychology. For instance, one particular view from linguistics is that the task for a psycholinguist is to hypothesize a language faculty separate from a general cognitive processing mechanism and test such a linguistic hypothesis through experimentation. Or one may hypothesize a particular grammar model with three linguistic domains, (a) linguistic primitives such as grammatical notions and features, including set relations (and possibly hierarchical relations of certain linguistic categories, though they can hold a particular set relation), (b) a computational system that deals with linearity (time) and memory, and (c) socio-cultural variations. These linguistic domains interact with each other, and a different grammatical theory arises in the model depending on how a certain grammatical construction or phenomenon is treated. A psycholinguist's role is to prove this kind of grammar model. Or an alternate view is to consider that all linguistic activities or behaviors are derived from general cognitive processes and that there is no need to theorize a specific linguistic faculty. That is, linguistic processing is not different from optimal information processing for the human brain (i.e., a biological architecture). A psycholinguist's role is to prove this kind of hypothesis. These are a few examples. Depending on the particular theoretical perspective a researcher subscribes to, his or her own specific investigative goals and interpretations differ. Ultimately, however, every psycholinguist's goal is to uncover mechanisms for language comprehension, production and acquisition processes.

All theoretical and experimental outcomes from such endeavors enrich our understanding of the language acquisition and processing mechanisms.[1] In the context of the current volume, the acquisition of the Japanese grammatical system in a broader sense and how the Japanese language is affected by a cognitive processing mechanism are discussed. Due to diverse approaches and interests in this field, psycholinguists often find themselves in serious theoretical debates. Japanese psycholinguists are, of course, no exception because Japanese is merely a variant of human language and because our perceptual and production mechanisms are biologically universal. Since we humans share the same biological mechanisms, theoretical constructs often have universal implications. What one finds in the general psycholinguistics field is also observed in the Japanese psycholinguistics field.

**1** It is important to point out that understanding how the processing mechanism operates helps us formulate a theory of grammar, assuming that the relation between the grammar and the processing mechanism is transparent. In this view, one could either say that the processing mechanism shapes grammar or that grammar shapes the processing mechanism, depending on one's point of view. Therefore, processing and acquisition are closely related.

# 3 Japanese psycholinguistics

The study of the Japanese language in psycholinguistics has advanced quite significantly in the last half century due to the progress in the study of cognition and brain mechanisms associated with language acquisition, use, and disorders, and in particular, because of technological developments in experimental techniques. Below, two subfields of Japanese language acquisition and processing are outlined.

## 3.1 Japanese language acquisition

The theory of generative grammar raised the fundamental question of "How is language acquired?" (e.g., Chomsky 1965, 1986), and sought explanatory power in its grammatical theory to account for the acquisition phenomena. One could argue that this theoretical framework made the single most influential contribution to the investigation of first language (L1) as well as second language (L2) acquisition among various linguistic theories. Generative grammarians have investigated various issues with different interpretations of the innate language faculty, Universal Grammar.

Roughly speaking, children's grammatical rule acquisition was investigated during the 1970s in the L1 acquisition field, but with the development of the Principles and Parameters approach (Chomsky 1981), language universals and particulars were differentiated and more issues related to the principles and the parametric values of Universal Grammar were investigated in the 1980s and later. While some innate principles were emphasized, a more traditional statistical learning approach, such as the usage-based theory (Tomasello 2003), had been advocated as well. The Universal Grammar approach did not deny statistical learning or a hypothesis testing approach in children's language acquisition (Young 2004), but it presented the "poverty of stimulus" argument and denied that statistical learning alone can explain grammar acquisition. See the debate between Stephen Crain and Michael Tomasello in the Boston University Language Development Conference (2004). In the Minimalist approach (Chomsky 1993, 1995), the perceptual/sensory component makes use of computational skills and does statistical learning. However, this learning alone still does not explain everything about language acquisition without assuming some primitives and structural relationships (Chomsky 2010). Recently, those who pursue this kind of Universal Grammar approach have been developing a biolinguistic program. On the other hand, those in the usage-based approach work more on computational exploration. This old nature vs. nurture debate has driven the development of theories of grammar acquisition.

Without question, the L1 Japanese acquisition field has been influenced by this general trend. Developmental psychologists were often concerned with vocabulary and construction acquisition, and more spontaneous speech data were reported in the earlier days (e.g., Okubo 1967, Noji 1974). The description of which word/

construction was acquired when received much attention before the 1980s (see Clancy 1985). With the advancement of linguistic analyses of the language, however, more issues related to Universal Grammar were investigated, as mentioned above, and Japanese made significant contributions to a general theory in this framework. Independently, investigations into pragmatics and sociocultural aspects have also been promoted, especially, with the influence of Schieffelin and Ochs (1986) because Japanese is quite different from languages like English. Today one can see vastly different issues being investigated across different linguistic subfields such as phonetics, phonology, morphology, syntax, semantics, pragmatics, discourse, sociolinguistics, neurolinguistics, and so on, though neurolinguistic studies on Japanese children are not as advanced as those on other languages due to technical and ethical reasons. Moreover, the advancement of computer technology has brought us opportunities to simulate language acquisition models, e.g., connectionist/computational learning models. We will see some of the above discussed studies in this volume.

The advancement of different linguistic theories also affected the field of adults' second language acquisition.[2] Before the 1980s more attention was paid to educational linguistics (and pedagogy), i.e., what is acquired and what is not, than to the brain function (or acquisition mechanism). Similar to the case of L1 acquisition, Chomsky's theory of Universal Grammar has had a profound influence on the development of L2 acquisition theories (White 1989, 2003). Many early L2 acquisition studies within the Principles and Parameters approach dealt with UG accessibility. That is, it examined whether Universal Grammar is (a) not accessible at all, (b) partially accessible (only those instantiated in L1) or (c) fully accessible during the course of L2 acquisition. Whether adults could access Universal Grammar became

**2** Of course, child second language acquisition or bilingualism is also influenced. However, we do not discuss bilingualism as it is one of the topics in the *Handbook of Applied Linguistics* (edited by M. Minami) in the HJLL series. Note that adult and child second language acquisition differ in the following aspects, as pointed out in the *Behavioral and Brain Sciences* special issue: (a) adults who are cognitively mature, unlike children, can use different learning strategies (e.g., production system, problem solving, complicated reasoning, etc.) (Bley-Vroman 1996, Li 1996, Otero 1996). (b) Adults can use L1 knowledge to learn a foreign language, i.e., L1 transfer (facilitating or impeding L2 acquisition) (Schachter 1974, 1988, Bhatt and Hancin-Bhatt 1996). (c) Adults can benefit from classroom instruction, textbooks, and learning tools in addition to data in the natural context (Li 1996, MacWhinney 1996). (d) Most adults do not reach the proficiency level that children can reach (Bickerton 1996) and adults have more individual variations in their acquisition level (Bley-Vroman 1996, Sorace 1996, Vainikka and Young-Scholten 1996). See morpho-syntax acquisition in this chapter. (e) Adults take more time to acquire L2 (Newmeyer 1996). (f) Children can acquire L2 without conscious efforts, whereas adults need fairly conscientious efforts (Freidin 1996, Li 1996). Because of these differences, it has been argued that it is important to separate the investigation of adult and child acquisition processes. However, these differences should not be critical in the study of Universal Grammar/biolinguistics because it is not an investigation of the computational system, but rather the biological endowment. See also Suzuki and Shirahata (2013).

an important issue, given the fact that L2 learners' ultimate attainment, i.e., L2 grammar, often does not become like that of native speakers, i.e., L1 grammar. Furthermore, the following questions were raised: What is the initial state of L2 grammar? Does it have L1 parametric values (Schwartz and Sprouse 1994) or the initial (or default) values in the grammar as in those during the course of L1 acquisition? These questions brought certain insightful outcomes in the parameter-resetting approach, but also there remained unexplained data. For a general overview, see White (2000).[3]

Many of the early L2 Japanese acquisition studies fell in applied linguistics and educational linguistics. However, again, generative grammarians investigated the above issues in L2 Japanese (although there were many more studies in L2 English by Japanese speaking learners). One of the methodological differences from L1 acquisition studies is that many L2 studies have employed written questionnaires because of convenience, i.e., the participants in their studies are often adults and almost everyone examined is literate (with a longer attention span than children), unless the issues investigated are relevant to spoken language. Today, as the number of L2 Japanese speakers increases, one finds more L2 Japanese studies regarding what is acquired when and why, as well as the acquisition of pragmatic and sociocultural aspects.

## 3.2 Japanese language processing

Japanese language processing is a much younger field than language acquisition. Among the subfields of language processing, however, orthographic and lexical processing has the longest history. Orthographic processing received attention in Japanese because of the orthographic difference from English type alphabetic processing. Mature Japanese readers decode three different Japanese scripts, *hiragana* (cursive *kana*), *katakana* (square *kana*), and *kanji* (Chinese characters). These three different characters are usually mixed within a single sentence, although in principle a sentence can be written in only *kana* (or *kanji*). In present-day Japanese, *hiragana* is primarily used to indicate high-frequency morphemes such as case markers, postpositions and inflectional endings while *katakana* is used for loan words except those of Chinese origin and for emphasis (e.g., onomatopoetic words and foreigners' conversations in comics). Since these script forms (46 basic *kana*: 71 with the use of diacritics, used to indicate voicing, for example) are moraic, their script-sound correspondence is highly regular. *Kanji*, on the other hand, do not have predictable, regular script-sound correspondences. They are primarily used for nouns and the roots of adjectives and verbs. Because of these different scripts, researchers were interested in questions such as whether different orthographies constitute different

3 For the recent interface approach, see White (2011).

processing routes. For instance, since *kana* has a regular script-sound correspondence, many scholars have proposed that it is likely to need phonological mediation (i.e., indirect access).[4] On the other hand, *kanji* can take the direct route because they do not have obvious regular script-sound correspondences (i.e., word meanings are directly retrieved from the visual representation of the words without phonological mediation).[5] For more on orthographic and lexical processing, see Nakayama (1999), Saito (2006) and Wydell (2006) and the references cited therein.

Investigation into Japanese sentence processing became more active after the late 80s. Many questions have been raised such as how a sentence is parsed, how lexical and syntactic ambiguities are resolved, how a gap is filled in a sentence, and why certain sentences are more difficult to process than others. These are also general questions one can raise during the investigation of any language. Their answers help us understand how the human brain functions, i.e., whether there is a universal processing mechanism in the human brain.

Understanding an adults' processing mechanism further relates to understanding the language acquisition mechanism. Research questions, such as what is the initial state of the processing mechanism and how do children use their parser when formulating their grammar, allow us to understand the initial state of the language faculty and how one's grammar develops. Furthermore, these questions can be addressed to language impaired learners as well. The comparison of data from both normal and language impaired learners provides us further insights into issues of how our language faculty functions. Thus, the processing field is closely connected with theoretical linguistics, language acquisition, neurolinguistics, computational linguistics, and other fields included in cognitive science, all of which open a door to the understanding of the human mind.

Studies on the Japanese language have contributed greatly to the field from a cross-linguistic perspective. For instance, it is a head-final language, in which the important verb information comes at the end of a verb phrase (and a sentence), unlike a head-initial language like English. Therefore, parsing models created based on English data alone could not accommodate languages like Japanese, and could not become universal sentence processing models. Which model, a serial or a parallel processing model, should be employed as a universal sentence processing model was one of the issues debated from the 1990s until recently. Constructing a universal processing model has so far been the predominant approach in the field, but there is no a priori reason that one cannot propose a language specific processing model.

---

**4** Impairment of *kana* processing has been reported in Sasanuma and Fujimura (1971) among aphasic subjects with the additional symptom of apraxia of speech.

**5** Phonological processing in *kanji* is also observed in Horodeck (1987), Wydell, Patterson, and Humphreys (1993) among others. For instance, Wydell, Patterson, and Humphreys claim that reading *kanji* is characterized by parallel access to semantics from orthographic and phonological representations.

See Nakayama (1999) and Miyamoto (2008) for overviews of various issues and parsing models. Today more research emphasis is on our brain functions, e.g., which part of the brain is activated when. However, we are still in the "fact finding" process and far from answering "Why?" questions.

The field of L2 Japanese processing is in its infancy. The issues studied are similar to those in L1 lexical and sentence processing. The number of factors involved in processing increases greatly in L2 compared to L1 because there exist (at least) two grammars and L2 grammar is not necessarily native-like. Furthermore, given limited real time experience, the processor is not always working as fast as it can optimally. These are additional considerations that make L2 processing studies difficult to carry out. As pointed out elsewhere in this Handbook, more psycholinguistic investigations are sought in this field.

## 4 This volume

This volume brings together state-of-the-art findings and discusses our brain functions, specifically, the process of Japanese language acquisition – how we acquire/learn the Japanese language as a first or second language – and the mechanism of Japanese language perception and production – how we comprehend and produce the Japanese language. In turn we address the limitations of our current understanding of the language acquisition process, and perception and production mechanism. Issues for future research on language acquisition and processing by users of the Japanese language will also be addressed.

The volume chapters are conventionally placed into two larger areas, language acquisition and language processing, but the contents of those chapters are by no means exclusive of each other as they are naturally interrelated. For instance, the language acquisition process involves language processing, and language processing depends on what is acquired. This volume is intended for experienced linguists or cognitive psychologists who are interested in cross-language differences and who would like to do a comparative study and can follow the importance of the issues under discussion and understand the latest analyses in Japanese psycholinguistics as well as Japanese language users with some knowledge of linguistics, psychology, and/or cognitive sciences (e.g., graduate students in those fields).

Some terminological notes are in order here. The term "language acquisition" is broadly used and is not distinguished from "learning" in a technical sense. It also includes first language (L1), second language (L2), third language (L3), and bilingualism. Furthermore, the term "second language acquisition" is used broadly including language development processes where Japanese is a second language or a foreign language, unless specifically stated so.

It is important to note that the editorial approach for this volume is slightly different from our previous handbook (Nakayama et al. 2006). Because the current handbook is one of the volumes in the HJLL series, and written in English with space limitations, it is not encyclopedic and is narrowly focused. For instance, topics related to educational linguistics and language pedagogy are not included because they are in the *Handbook of Japanese Applied Linguistics* (edited by Masahiko Minami) in the HJLL series. Our coverage crucially differs from the *Applied Linguistics* volume in that this volume looks at the acquisition of grammatical knowledge and the development of our cognitive systems, and it does not necessarily provide chapters discussing the Japanese language from sociocultural and educational perspectives. Furthermore, the selected topics are primarily of current theoretical interest and those that have demonstrated significant research outcomes within the past five to ten years, i.e., since the publication of Nakayama et al. (2006). The selected topics in this volume have promising theoretical contributions to other human languages in the fields of First Language Acquisition, Second Language Acquisition, and Language Processing.

Some topics that were not included are script processing, lexical processing, and discourse processing. L1 character/lexical processing has been discussed extensively elsewhere (e.g., see chapters on this topic in Nakayama et al. 2006). Although Sakamoto's chapter discusses pragmatic and contextual coherence, the language processing section only infrequently referred to discourse processing. This is an area that still awaits more investigation, but some promising results have been found in Hirotani and Schumacher (2011) and Wang and Schumacher (2013), where contextual relevance is discussed. On the other hand, despite fewer published works in Japanese, Specific Language Impairment (SLI) was selected because of the contributions Japanese SLI can make to recent cross-linguistic discussions of this topic. Also included in this volume is the acquisition of L2 English by Japanese adults and children because it discusses how L1 Japanese has affected L2 English during the course of its acquisition in mature brains and in developing brains. In addition, many chapters reflect current popular methodologies such as the truth value judgment task, eye tracking, Event Related brain Potentials (ERP), and functional Magnetic Resonance Imaging (fMRI). Despite these specific editorial choices, and with this focused approach, the Handbook presents significant findings that Japanese psycholinguistic studies offer to various theories of the human mind. The variety of topics dealt with in this volume indicates the scope of scholarly inquiries on the Japanese language and the maturity of the field.

I will now briefly explain the subfields in relation to the topics selected in this volume.

## 4.1 First language acquisition

Six topics were selected in the first language acquisition section. These chapters demonstrate an advancement of L1 language acquisition studies in Japanese. They

allow us to understand how perceptual, conceptual, and grammatical features (including pragmatic and discourse features) are developed and how our brain and perceptual system develop during the course of Japanese language acquisition, specifically from both language particular and universal perspectives.

First, Mazuka's chapter "Learning to become a native listener of Japanese" discusses how a baby becomes a Japanese listener. An infant's perceptual system is attuned to the L1 phonological system within the first year of life. It has not been possible to investigate this kind of infant's perceptual development before, but with technological developments, she and her colleagues have made it possible. Since Japanese segmental and suprasegmental characteristics differ from those of well studied European languages such as English, the findings make a significant contribution to theories of a child's perceptual development and phonological development in grammar formation.

Imai and Kanero's chapter "The nature of the count/mass distinction in Japanese" discusses one of the most fundamental conceptual distinctions. Since Japanese does not morphologically mark the singular and plural distinction on individuated object nouns all the time as English does, one wonders how these concepts manifest in Japanese children's grammar. That is, how do children make the ontological distinction (object vs. substance) even when their language does not have apparent morphosyntactic marking of count nouns and mass nouns? This is addressed in the chapter.

Fukuda, Fukuda, and Ito's chapter "Grammatical deficits in Japanese children with Specific Language Impairment (SLI)" discusses the grammatical deficits Japanese SLI children exhibit. It shows the contribution Japanese SLI can make toward a cross-linguistic theory of SLI. Despite fewer published works in Japanese, SLI was selected as a topic because of the contributions Japanese can make to recent cross-linguistic discussions of this topic. The chapter also offers a window to understanding how our brains develop and become efficient language users. In addition, this chapter is related to Murasugi's chapter "Root infinitive (RI) analogues in child Japanese" because of the discussion on the use of tense and aspect morphemes. Languages such as English clearly observe a developmental stage where children do not seem to produce utterances with tense. However, Japanese children do not seem to have such a stage because their verb forms always have tense morphemes. Murasugi challenges this view and shows that Japanese is no different from other languages in that Japanese children also have a developmental stage without tense, i.e., an RI analogue stage, but the stage occurs earlier than RIs in European languages. These two chapters by Fukuda, Fukuda, and Ito, and Murasugi make a fascinating theoretical contribution to cross-linguistic analyses.

Goro's chapter "The acquisition of constraints on quantifier scope" looks at Japanese children's scope interpretations, which are assumed to be the same as the adults' given the lack of clear negative evidence. However, empirical data suggest otherwise, i.e., similar to English. He first explains language-specific constraints on

scope interpretation from the learnability perspective, then reviews the data from recent experimental studies, and discusses the consequences of the findings.

How do young Japanese children develop narrative structures? How do adults/parents guide their children in the acquisition of culturally appropriate styles of narrative and literacy? These questions are explored in Minami's chapter "Narrative development in L1 Japanese". It specifically examines children's narrative discourse styles and the role of parental input in facilitating the development of children's personal narratives. It also discusses the relationship between sociocultural background and the development of literacy in young children.

Although Chang's chapter on Japanese processing is not included in the language acquisition section, it is highly relevant to the discussion of language acquisition. His model shows successful statistical learning as well as issues this kind of model poses. The statistical learning model is relevant to Tomasello's (2003) usage-based language acquisition model, among others, which is often contrasted with nativist views such as the principles and parameters framework and the minimalist framework of generative grammar.

## 4.2 Second language acquisition

Four chapters discuss issues in second language acquisition. All chapters refer to the influences of L1. L2 acquisition is crucially different from L1 acquisition because one grammar already exists in a language user's brain during the course of L2 acquisition. Thus, L2 language development or the process of learning L2 Japanese (or L2 English) is different from L1 acquisition. Similarities and dissimilarities between L1 and L2 allow us to lay out similar and dissimilar brain functions in language acquisition.

Shirai's chapter "The L2 acquisition of Japanese" sets the stage with a summary of previous studies in L2 acquisition research. Taking a different approach from Mori and Mori (2011), which reviewed research on L2 acquisition and instruction of Japanese from 2001 to 2010, it briefly summarizes frequently cited studies in L2 acquisition research. Few psycholinguistic topics seen in this volume appeared in this summary. It is because they are more related to pedagogy and applied linguistics, which is easily understood when one considers who wrote and referred to those L2 Japanese acquisition articles in English. Theoretical papers were published more often in Japan and they were frequently about L2 English acquisition. Thus, the chapter also provides other important studies that did not necessarily appear in the citation index, and points out their significant findings and issues for future research.

Adult L2 learners often make errors. Some of them are similar to the errors children make during their L1 acquisition whereas others are different from those of L1 children. Nakayama and Yoshimura's chapter "The modularity of grammar in L2 acquisition" presents a theoretical framework that accounts for the complexity

of errors L2 learners make including fossilized errors in L2 English and Japanese. Assuming different grammatical modules such as phonological, morphological, syntactic, semantic, and pragmatic components, they claim that more errors persist if they fall into more than one grammatical component. Furthermore, added are performance factors (e.g., computational limitation) that bring more complications during the course of language acquisition. This interface approach can account for the comprehension/production processes more readily than the earlier framework, the Principles and Parameters approach.

Telicity means the aspectual property of a verb phrase that indicates that an action or event has a clear endpoint. The semantic notion telicity exists universally in human languages, but the grammatical manifestation of this notion differs depending on the language. Gabriele and Hughes' chapter "Tense and aspect in Japanese as a second language" deals with this issue. It first reviews the relevant linguistic facts on tense and aspect in Japanese, and then discusses several recent studies that have examined the Aspect Hypothesis (Andersen and Shirai 1996) and L1 transfer in L2 Japanese at the levels of grammatical aspect, lexical aspect, and within a noun phrase. In addition, the processing of tense and aspect in L2 Japanese are also discussed.

How do language acquisition and brain development co-occur in early childhood? When does cortical plasticity for the language modules deteriorate? Hagiwara's chapter "Language acquisition and brain development: cortical processing of a foreign language" addresses these questions. Thanks to great progress in the field of cognitive neuroscience over the past decade, new imaging techniques allow us to answer a long debated question, whether complete mastery of a language is impossible after puberty (i.e., after the critical period). By examining the status of L2 English in the Japanese brain both in adulthood and childhood, it shows that a certain aspect of syntax, core computation in narrow syntax and morphology-syntax interface, is free from the notion of the critical period, and that lexical learning in childhood is biologically constrained in the human brain. These findings make a significant contribution to a general theory of L2 acquisition. See also Mori and Calder (2013) on Japanese sojourners/expatriates' L1 Japanese and L2 English vocabulary acquisition.

### 4.3 L1 Japanese processing

The L1 Japanese processing section consists of five chapters on sentence processing. The selected topics included in the volume are rather limited, due to the relative youth of the field, and in order to avoid overlap with the topics covered in Nakayama et al. (2006). However, these chapters allow us to understand how the Japanese language is processed in the native speakers' cognitive systems and the findings bring implications to language particulars and universals. Although the data discussed in first language acquisition chapters (except the SLI chapter) come from

the spoken language, the data in the sentence processing chapters are rather mixed. This difference arises for a historical reason, i.e., the methodological ease of presentation of the written language and recent technological developments in testing the spoken language.

Hirose's chapter "Resolution of branching ambiguity in speech" considers the role of prosody in distinguishing the two alternative structures, e.g., *midori no inko no mafuraa* 'a scarf with a green parrot' vs. 'a green scarf with a parrot'. The role of prosody in production and comprehension of NPs with a branching ambiguity is less clear and the relationship between syntactic structure, prosodic structure, and prosodic realization depends on different factors such as phonological and discourse factors.

Chang's chapter "The role of learning in theories of English and Japanese sentence processing" presents a statistical learning model based on a connectionist approach, i.e., different layers of processing. This model learns language-specific syntactic representations and uses these representations in incremental sentence production. It offers a more parsimonious account of psycholinguistic phenomena, where "universal" processing biases arise from language-specific knowledge. It addresses how frequency shapes the preference of a particular word order and what to anticipate in Japanese. In addition, as mentioned above, this chapter discusses implications of the model for language acquisition.

Koizumi's chapter "Experimental syntax: word order in sentence processing" also discusses word order since a flexible word order could increase processing complexity. By examining word order, the chapter illustrates the types of experimental studies currently underway in the field of syntax, e.g., evaluation of processing and linguistic theories by testing their predictions. The methodology and the argument used to decide Japanese base word order is also applied to another language, Kaqchikel (spoken in Central Guatemala), whose word order is either SVO or VOS. The chapter exemplifies a successful contribution to cross-linguistic studies stemming from the Japanese sentence processing field.

One of the most complex and well-studied grammatical structures is the relative clause structure. Kahraman and Sakai's chapter "Relative clause processing in Japanese psycholinguistic investigation into typological differences" discusses various factors involved in processing Japanese relative clause sentences. They evaluate the theories on filler-gap dependency formation such as Dependency Locality Theory (Gibson 2000) and Structural Distance Hypothesis (O'Grady 1997) in Japanese, and further look at other influential factors in processing such as frequency. The limitations and the possibilities of future directions in relative clause processing studies are also suggested.

Sakamoto's chapter "Processing syntactic and semantic information in the human brain: evidence from ERP studies in Japanese" examines the physiological evidence for dissociating syntactic and semantic processes by referring to ERP data. Based on this examination, it also evaluates two ("syntax-first" and "interactive")

parsing models that attempt to account for various types of linguistic processes. The chapter concludes that the Japanese processing mechanism works in an "expectancy-driven" way. That is, Japanese speakers do not wait to process a sentence until the end of the sentence, i.e., the verb, but rather they process it incrementally and anticipate forthcoming words.

## 4.4 L2 Japanese processing

The number of studies on L2 Japanese processing is quite small because it is difficult to conduct a study with a large population with a similar L1 background and proficiency levels, unlike L2 English. Experimental research designs and methodologies employed in the studies are similar to those in the first language. However, factors one must consider in experimental designs seem to be similar to those of L2 acquisition studies such as limited vocabulary, limited *kanji* knowledge, familiarity range of the voice quality in stimuli, articulatory (motor control) issues and so on. The L2 processing section contains three chapters whose topics are rather distinct.

Sawasaki and Kashiwagi-Wood's chapter "Issues in L2 Japanese sentence processing: similarities/differences with L1 and individual differences in working memory" considers whether L2 sentence processing is similar to or different from L1 sentence processing. As in Kahraman and Sakai's chapter, it discusses relative clause structures by evaluating the Dependency Locality Theory and the Structural Distance Hypothesis. Since working memory is believed to influence relative clause processing and explain individual differences in processing and comprehension performance, it also reviews research on this topic in L2 Japanese.

There are few studies on Japanese sentence production even in L1. L2 oral production studies are quite scarce including lexical pronunciation and accent. Iwasaki's chapter "Sentence production models to consider for L2 Japanese sentence production research" discusses both L1 and L2 Japanese sentence production in specific theoretical models. By referring to grammatical encoding, it points out differences between sentence production processes in European languages and Japanese. It also discusses the implications L2 (bilingual) sentence production models would bring to L2 Japanese, which have been created primarily based on L1 and L2 European languages.

Tamaoka's chapter "Processing of the Japanese language by native Chinese speakers" points out issues created by Chinese speaking L2 Japanese users. Because of their knowledge of Chinese characters or *kanji*, L1 Chinese readers pose different issues from those of English speaking L2 Japanese users, for instance, in their L2 reading. The chapter also reviews studies on lexical pitch accent and morphosyntactic processing. Because many of these studies cited in this chapter are published in Japanese, English speaking readers may gain new insights that shed some light on the language processing mechanism in general.

# 5 Concluding remarks

This chapter provided a brief summary of the theoretical trends in the Japanese psycholinguistics field. Among the many psycholinguistic issues investigated, the Handbook specifically focuses on Japanese language acquisition and processing mechanisms, mostly, at the phrase and sentence levels. Although not all topics were covered in this volume, what is delivered here is a useful collection of representative theoretical contributions from the Japanese language to studies of other human languages in the world. The studies discussed in the chapters in this volume appeal to both language particular and universal perspectives, as they advance our understanding of how the human brain works, and will attract an international audience who is interested in interdisciplinary language science studies. I hope the readers will learn as much from this Handbook as I did while editing it.

# Acknowledgments

I would like to thank Masahiko Minami, Jennifer Nakayama, Shinsuke Tsuchiya, and Noriko Yoshimura for their helpful comments on the earlier version of this chapter.

# References

Anderson, Roger W. and Yasuhiro Shirai. 1996. Primacy of aspect in first and second language acquisition: The pidgin/creole connection. In William C. Ritchie and Tej K. Bhatia (eds.), *Handbook of second language acquisition*, 527–570. San Diego, CA: Academic Press.

Bhatt, Rakesh M. and Barbara Hancin-Bhatt. 1996. Transfer in L2 grammars. *Behavioral and Brain Sciences* 19(4). 715–716.

Bickerton, Derek. 1996. A dim monocular view of Universal-Grammar access. *Behavioral and Brain Sciences* 19(4). 716–717.

Bley-Vroman, Robert. 1996. What we have to explain in foreign language learning. *Behavioral and Brain Sciences* 19(4). 718.

Boston University Conference on Language Development. 2004. Where does grammar come from? A debate on the nature of child language acquisition with Stephen Crain and Michael Tomasello, November 6, 2004. DVD. Boston: Boston University Conference on Language Development.

Chomsky, Noam. 1965. *Aspect of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of language*. New York. Praeger Publishers.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Ken Hale and Samuel J. Keyser (eds.), *The view from Building 20*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2010. Poverty of stimulus: Some unfinished business. Lecture presented at Conférence scientifique de Noam Chomsky du 29 mai 2010. Uploaded on October 24, 2012, at http://www.youtube.com/watch?v=OSFgTuHQyvo.

Clancy, Patricia. 1985. Acquisition of Japanese. In Dan Isaac Slobin (ed.), *The crosslinguistic study of language acquisition, vol. 1, Data*, 373–524. Hillsdale, NJ: Lawrence Erlbaum Associates.

Freidin, Robert. 1996. Adult language acquisition and Universal Grammar. *Behavioral and Brain Sciences* 19(4). 725–726.

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Alec P. Marantz, Yasushi Miyashita and Wayne O'Neil (eds.), *Image, language, brain: Papers from the first mind articulation project symposium*, 95–126. Cambridge, MA: MIT Press.

Herrmann, Theo. 2004. Was ist eigentlich Sprachpsychologie? *Journal of Foreign Language Education and Research* 7. 79–93.

Hirotani, Masako and Petra B. Schumacher. 2011. Context and topic marking affect distinct processes during discourse comprehension in Japanese. *Journal of Neurolinguistics*, 24(3), 276–292.

Horodeck, Richard A. 1987. The role of sound in reading and writing kanji. Ithaca, NY: Cornell University dissertation.

Li, Ping. 1996. Why don't L2 learners end up with uniform and perfect linguistic competence? *Behavioral and Brain Sciences* 19(4). 733–734.

MacWhinney, Brian. 1996. Language is learned. *Behavioral and Brain Sciences* 19(4). 735–736.

Minami, Masahiko (ed.). Forthcoming. *Handbook of Japanese Applied Linguistics*. Boston: De Gruyter Mouton.

Miyagawa, Shigeru and Mamoru Saito (eds.). 2008. *The Oxford handbook of Japanese linguistics*. New York: Oxford University Press.

Miyamoto, Edson T. 2008. Processing sentences in Japanese. In Shigeru Miyagawa and Mamoru Saito (eds.), *The Oxford handbook of Japanese linguistics*, 217–249. New York: Oxford University Press.

Mori, Yoshiko and Toshiko M. Calder. 2013. Bilingual vocabulary knowledge and arrival age among Japanese heritage language students at *hoshuukoo*. *Foreign Language Annals* 46. 290–310.

Mori, Yoshiko and Junko Mori. 2011. Review of recent research (2000–2010) on learning and instruction with specific reference to L2 Japanese. *Language Teaching* 44(4). 447–484.

Nakayama, Mineharu. 1999. Sentence processing. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 398–424. Boston: Blackwell.

Nakayama, Mineharu, Reiko Mazuka and Yasuhiro Shirai (eds.). 2006. *East Asian psycholinguistics 2: Japanese*. Cambridge, UK: Cambridge University Press.

Nakayama, Mineharu, Reiko Mazuka and Yasuhiro Shirai (eds.). 2006. *Handbook of East Asian psycholinguistics, Vol. 2: Japanese*. Cambridge, UK: Cambridge University Press.

Newmeyer, Fredirick J. 1996. Some incorrect implications of the full-access hypothesis. *Behavioral and Brain Sciences* 19(4). 736–737.

Noji, Junya. 1974. *Yōjiki-no gengo seikatu-no jittai* [Realities of language life in childhood]. Hiroshima: Bunka hyōronsha.

O'Grady, William. 1997. *Syntactic development.* Chicago, IL: University of Chicago Press.

Okubo, Ai. 1967. *Yōji gengo-no hattatsu* [Child language development]. Tokyo: Tokyo-do.

Otero, Carlos P. 1996. Language growth after puberty? *Behavioral and Brain Sciences* 19(4). 738–739.

Saito, Hirofumi. 2006. Orthographic processing. In Mineharu Nakayama, Reiko Mazuka and Yasuhiro Shirai (eds.), *East Asian psycholinguistics 2: Japanese*. 233–240. Cambridge, UK: Cambridge University Press.

Sasanuma, Sumiko and Osamu Fujimura. 1971. Selective impairment of phonetic and non-phonetic transcription of words in Japanese aphasic patients: Kana vs. kanji in visual recognition and writing. *Cortex* 7(1). 1–18.

Schachter, Jacqueline. 1974. An error in error analysis. *Language Learning* 24. 205–214.

Schachter, Jacqueline. 1988. Second language acquisition and its relationship to Universal Grammar. *Applied Linguistics* 9(3). 219–235.

Schieffelin, Bambi and Elinor Ochs. 1986. *Language socialization across cultures*. New York: Cambridge University Press.

Schwartz, Bonnie and Rex Sprouse. 1994. Word order and nominative case in non-native language acquisition: A longitudinal study of German interlanguage. In Teun Hoekstra and Bonnie Schwartz (eds.), *Language acquisition studies in generative grammar*, 317–368. Amsterdam: John Benjamins.

Sorace, Antonella. 1996. On gradience and optionality in non-native grammars. *Behavioral and Brain Sciences* 19(4). 741–742.

Suzuki, Takaaki and Tomohiko Shirahata. 2013. *Kotoba-no shūtoku* [Language acquisition]. Tokyo: Kurosio Publishers.

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tsujimura, Natsuko (ed.). 1999. *The handbook of Japanese linguistics*. Boston: Blackwell.

Vainikka, Anne and Martha Young-Scholten. 1996. Partial transfer, not partial access. *Behavioral and Brain Sciences* 19(4). 744–745.

Wang, Luming and Petra B. Schumacher. 2013. New is not always costly: Evidence from online processing of topic and contrast in Japanese. *Frontiers in Psychology* 4(363). [doi: 10.3389/fpsyg.2013.00363]

White, Lydia. 1989. *Universal Grammar and second language acquisition*. Philadelphia: John Benjamins.

White, Lydia. 2000. Second language acquisition: From initial to final state. In John Archibald (ed.), *Second language acquisition and linguistic theory*, 130–155. Malden, MA: Blackwell.

White, Lydia. 2003. *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

White, Lydia. 2011. Second language acquisition at the interfaces. *Lingua* 121. 577–590.

Wydell, Taeko N. 2006. Lexical access. In Mineharu Nakayama, Reiko Mazuka and Yasuhiro Shirai (eds.), *East Asian Psycholinguistics 2: Japanese*. 241–248. Cambridge, UK: Cambridge University Press.

Wydell, Taeko N., Karalyn E. Patterson and Glyn W. Humphreys. 1993. Phonologically mediated access to meaning for kanji: Is a ROWS still a ROSE in Japanese kanji? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19(3). 491–514.

Yang, Charles D. 2004. Universal Grammar, statistics or both? *Trends in Cognitive Sciences* 8(10). 451–456.

# I Japanese Language Acquisition

Reiko Mazuka

# 1 Learning to become a native listener of Japanese

## 1 Introduction

Infants are born with an ability to learn any human language. But within the first year of their lives, their perceptual systems become attuned to the phonological system of the ambient language, a development which is a prerequisite for language acquisition. To date however, research on how infants learn the phonological system of a language has been carried out almost exclusively on the basis of English and several other European languages such as Spanish, Italian, French, German, and Dutch. In comparison, research on the development of non-European languages is severely lacking. Japanese is a language whose segmental and suprasegmental characteristics differ from those of European languages in critical ways, such as mora-timed rhythm, edge prominent prosody, presence of duration-based phonemic contrasts and exceptional distribution of segments. By taking advantage of language-specific properties of Japanese, we can investigate how a particular language structure interacts with the human cognitive system such that human children can learn to process language with seeming effortlessness. In the present chapter, we will discuss how research on Japanese contributes to important questions on infants' phonological development by introducing three lines of research of our own: studies on phonemic segment acquisition, lexical level prosody, and phonological grammar development.

## 2 Duration-based phonemic contrasts

The first set of studies investigated how infants acquire the phonemic categories of their native language. During the past several decades, research on infant speech perception has demonstrated that the ability of infants to discriminate speech segments undergoes a significant shift during the first year of life, from sensitivity to a broad range of contrasts to becoming attuned to their native language (e.g., Kuhl 2004; Saffran, Werker and Werner 2006; Werker and Yeung 2005, for review). It appears that infants start out with the sensitivity to discriminate among a majority of segmental contrasts, including contrasts that are not used in their native languages (Eimas et al. 1971; Kuhl et al. 2006; Polka and Werker 1994; Tsushima et al. 1994; Werker et al. 1981). The ability to discriminate nonnative speech sounds begins to decline or disappear by 6 months for vowels and by 10 months for consonants (Best and McRoberts 2003; Kuhl et al. 1992, 2006; Polka and Werker 1994; Tsushima

et al. 1994; Werker and Tees 1984). There are, however, other developmental paths possible. Some acoustically quite distinct contrasts, such as those between click sounds, remain discriminable even without any experience in one's native language (Best, McRoberts and Sithole 1988). A third pattern involves infants' experience with a particular language facilitating or enhancing their ability to discriminate the contrasts. This pattern has been reported, for example, for the contrast between voiced and voiceless stop consonants in Spanish (Eilers, Wilson and Moore 1979; Lasky, Syrdal-Lasky and Klein 1975), the alveolar versus dental contrast (/d/ vs. /ð/) in English (compared to French infants) (Polka, Colantonio and Sundara 2001; Sundara, Polka and Genesee 2006), affricate-fricative contrasts in Mandarin (Tsao, Liu and Kuhl 2006), a nasal phonetic contrast (/n/ vs. /ng/) (Narayan 2006), and the /r/ and /l/ contrast in English (Kuhl et al. 2006; Kuhl et al. 2001).

In most languages, the majority of phonemic categories are primarily characterized by acoustic quality (spectral) changes that make up consonants and vowels. Consequently, studies on infants' sensitivity to phonemic contrasts to date have also focused on their ability to discriminate segments on the basis of "quality" differences. In addition to quality-based categories, some languages utilize phonemic categories that are based on quantity changes. Yet, Japanese and a small number of other languages use *duration*-based phonemic contrasts such as vowel duration (e.g., *toko* 'bed' vs. *tokoo* 'travel') and geminate obstruents (e.g., *tokoo* 'travel' vs. *tokkoo* 'thought police').

The perception of sound duration change itself is a general and primitive auditory function. It is reported that even guinea pigs exhibit responses for tone duration discrimination (Okazaki et al. 2006) and event related potential (ERP) studies have exhibited human newborns' responses to sound duration changes (Kushnerenko et al. 2001; Leppänen et al. 1999). Therefore, we assume that Japanese infants should also have sensitivity to the durational differences of auditory stimuli at a general level. It is not clear, however, how a sensitivity to vowel duration can develop in a linguistic context. Detecting isolated acoustic cues that differentiate phonemic contrasts does not necessarily mean the same cue in a phonemic context is equally discriminable. For example, in their classic study, Miyawaki et al. (1975) demonstrated that Japanese adults, who show great difficulty discriminating /r/ from /l/ in a linguistic context, are as good as English speaking adults in discriminating the acoustic cues that differentiate the two sounds when presented in isolation. One of our questions is whether infants learn duration differences by the same process as quality differences.

## 2.1 Japanese infants' discrimination of phonemic vowel-duration

We will discuss three studies on this topic. In the first study, we examined the developmental changes in Japanese infants' abilities to discriminate long and short

vowels. The acquisition of vowel duration contrasts presents infants with an extra challenge not typically found with quality-based contrasts: infants must learn to distinguish phonemic and non-phonemic use of the same acoustic cues. Variations in vowel duration can appear in any language, but how they are utilized phonemically differs from one language to another. Thus, in order for infants to learn the phonemic use of vowel duration contrasts, they must also be able to isolate them from other durational modulations that are not phonemic. These include factors such as intrinsic differences in vowel duration for high and low vowels (House and Fairbanks 1953; Peterson and Lehiste 1960), linguistic stress (Fry 1955; Oller 1973), contrastive stress and semantic novelty (Umeda 1975), final lengthening (Beckman and Edwards 1990; Cooper and Paccia-Cooper 1980) and speech rate (Hirata 2004).

How each of these factors influences vowel duration in a specific context will vary from language to language (Hoequist 1983; Vihman 1992). But in a small group of languages, including Japanese, Finnish, and Estonian, the distinction between long and short vowels is phonemic and primarily based on duration. A phonemically long vowel is not acoustically distinguishable from a vowel that has been made long due to phrasal modulation. Therefore, infants learning these languages must notice not only that some vowels in their input are regularly longer than others, but also that some of the vowels could be longer due to their position in a phrase (e.g., final lengthening) while others are lengthened independently of their position. Conflations of the same acoustic cues in phonemic and non-phonemic contexts may not be unique to vowel duration, but since the lengthening of vowels is one of the primary cues used for prosodic level phrasing, acquisition of phonemic categories defined by duration change must be one of the most challenging tasks an infant faces. Among the languages that use vowel duration contrasts phonemically, Japanese has an additional function not shared by Finnish or Estonian. In Japanese, the distinction between a short and a long vowel corresponds to one mora versus two moras, and is thus critical for the mora-timed rhythm of Japanese (Ladefoged 1975; Port, Dalby and O'Dell 1987).

To study how Japanese infants develop the ability to discriminate long and short vowels, Sato, Sogabe and Mazuka (2010a) tested 4-, 7.5- and 9.5-month old Japanese infants with a pair of nonsense words /mana/ vs. /ma:na/, in which the long-short vowel contrast is embedded in the first syllable of each bisyllabic word. An experimental procedure known as the Modified Visual Habituation paradigm was used (Stager and Werker 1997). In this paradigm, an infant is seated on his or her parent's lap in a dimly lit room while a computer monitor shows a checkerboard pattern over the entire screen. The infant is presented with a particular stimulus repeatedly. A video camera records the infant's face. An experimenter monitors the infant silently via a video monitor in the control room and holds down a key on a computer keyboard whenever the infant is looking straight at the monitor in the experiment room. The computer thus records the infant's looking time. It has been well established

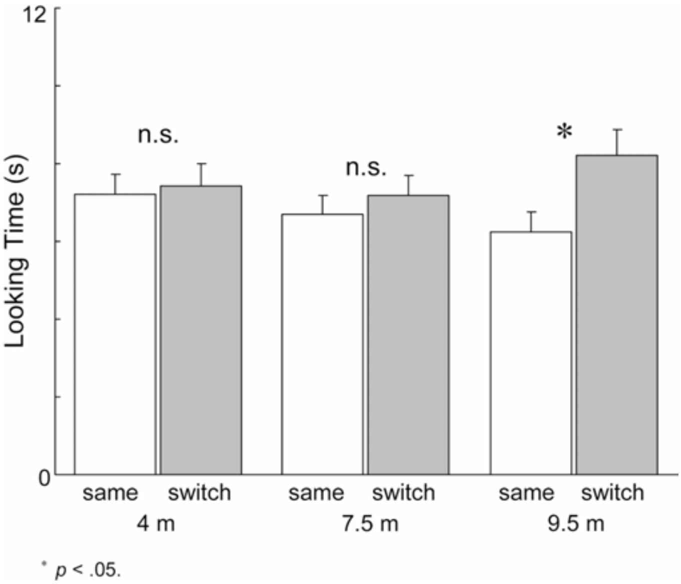



**Figure 1:** Mean looking times (with standard error bars) during the same and switch trials (see procedure) for the long/short vowel contrast in 4-, 7.5- and 9.5-month old groups (Experiments 1–3). Only the 9.5-month old group exhibited a significant difference in looking times between same and switch conditions (*: $p < .05$). (Sato, Sogabe and Mazuka 2010a: 111, Figure 2)

from previous studies that infants tend to stare at the screen when they pay attention to the auditory stimuli. As the novelty wears off they become bored and restless, looking less at the screen and spending more time looking around the room, at their own hands and feet, and so on. When the infant's looking time has declined to a predetermined level, such as 60% of the initial looking time, it is determined that the infant has habituated to the stimuli and the test trials begin. During the test trials, infants are presented with the same stimuli (same trial) and the other stimuli (switch trial) in a counter-balanced order, and their looking time between the two test trials are compared. If infants notice the difference between the habituated stimuli and the new stimuli, it is expected that their attention to the new stimuli recovers (in other words, looking time during the switch trials is longer than looking time during the same trial). If they do not discriminate between the two, there is no reason for their attention to recover, and the looking time between the same and the switch trials should not differ.

As Figure 1 demonstrates, Japanese infants were unable to discriminate this contrast at 4 months. They become capable of discriminating between the stimuli by 9.5 months of age. This is an enhancement pattern of development, whereby infants' poor discrimination at a younger age improves as they become older. As discussed above, this is atypical of how infants develop the ability to discriminate phonemic

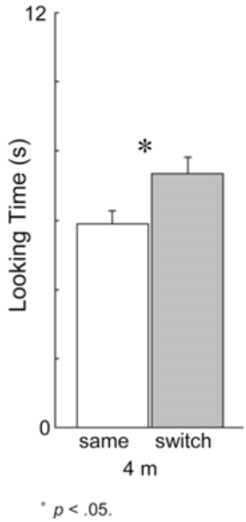



**Figure 2:** Mean looking times (with standard error bars) during the same and switch trials (see procedure) for the quality contrast (/mana/ vs. /mina/) in 4-month old infants (Experiment 5). The looking time during the switch trial was significantly longer than that during the same trial (*: $p < .05$). (Sato, Sogabe and Mazuka 2010a: 114, Figure 5)

contrasts (early discrimination ability declines with age in the absence of exposure to such contrasts in the native language). For comparison, Sato, Sogabe and Mazuka (2010a) also tested Japanese infants' ability to discriminate quality-based vowel contrasts in the same context, i.e., in /mana/ versus /mina/. It was found that 4-month old Japanese infants were able to discriminate the quality-based vowel contrast between /a/ versus /i/, as shown in Figure 2. These data suggest that the discrimination of duration-based vowel contrasts develops differently from that of quality-based vowel contrasts.

## 2.2 Discrimination of duration-based consonant contrasts

Durational differences in consonants are also used phonemically in languages such as Japanese (Vance 1987), Italian (Pickett, Blumstein and Burton 1999), Persian (Hansen 2004), Bengali and Turkish (Hankamer, Lahiri and Koreman 1989; Lahiri and Hankamer 1988). These contrasts are often referred to as single/geminate distinction. When stop consonants are involved, the difference between a single and a geminate consonant is characterized as the durational change in the closure, and the geminate stops have been reported to have one and a half to three times the acoustic closure duration of the single stops (Ladefoged and Maddieson 1996). It has been proposed that in some languages, such as Estonian and Sami, three

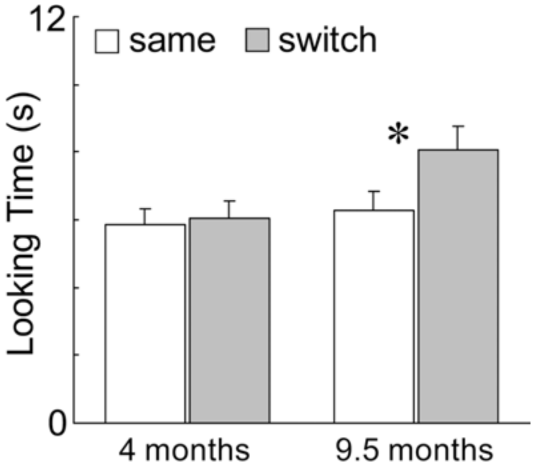



**Figure 3:** Mean looking times (with standard error bars) during the same and switch trials (see procedure) for naturally uttered /pata/ and /patta/ in 4- and 9.5-month-old infants (RC condition, Experiment 1). The 9.5-month-olds exhibited a significant difference in looking times between same and switch trials (*$p < .05$). (Sato, Kato and Mazuka 2012: 24, Figure 1)

distinctive lengths for consonants are distinguished (Engstand 1987; Lehiste 1966; Eek 1984–5). As with vowel duration contrasts, infants must learn whether or not their language utilizes the durational changes phonemically, and if it does, which of the durational changes they hear are phonemic and which are due to prosodic or other factors.

Sato, Kato and Mazuka (2012) examined how Japanese infants develop the ability to discriminate duration-based consonant contrasts. If the enhancement pattern of development they found in vowel-duration discrimination was indeed due to the fact that duration-based phonemic discrimination is difficult for young infants, a similar pattern of development should also be observed for the duration-based consonant discrimination. However, vowels and consonants differ not only in their acoustic properties but also in the roles they play in the phonology of a language. It is possible that the discrimination of vowels and of consonants follow different developmental trajectories. To test these predictions, Sato, Kato and Mazuka tested 4- and 9.5-month old Japanese infants using the same experimental paradigm (visual-habituation). Infants were presented with either /pata/ or /patta/ repeatedly until they habituated to the stimuli, and were then tested to ascertain whether or not they dishabituated when presented with the other stimuli. Figure 3 shows the results of these experiments. As was the case with vowel-duration contrasts, Japanese infants were unable to discriminate the contrast at 4 months of age but became able to discriminate them by 9.5 months of age.

The results of these studies showed that the ability of Japanese infants to discriminate duration-based phonemic contrasts develops in the enhancement pattern; they are not able to discriminate at an earlier age but become able to discriminate as they grow older. The fact that the same pattern of development was found for both vowels and consonants, and that Japanese infants were able to discriminate the quality-based vowel contrast /mana/ versus /mina/ at as early as 4 months of age, suggests that it is the duration-based nature of these contrasts that leads to the enhancement pattern.

A number of factors have been considered to account for why the duration-based phonemic contrasts are difficult for young infants to discriminate, while the quality-based contrasts seem to be readily discriminable. One set of factors that have been considered in Sato, Kato and Mazuka (2012) are the limited cognitive capacities of younger infants. Previous studies of infants' ability to discriminate auditory duration contrasts have indicated that perception of duration differences is ratio-based (VanMarle and Wynn 2006) and that the precision of discrimination improves between 6 and 12 months of age (Brannon, Suanda and Libertus 2007). Infants at 6 months of age are capable of discriminating durational differences with a ratio of 1:2, but not those with a ratio of 2:3. By 12 months, they become capable of discriminating 2:3 ratios as well. In Sato, Sogabe and Mazuka (2010a), the durational differences of vowels themselves were approximately 100 milliseconds versus 200 milliseconds. Similarly, the closure durations of single and geminate stops in Sato, Kato and Mazuka (2012) were approximately 100 milliseconds and 200 milliseconds, again demonstrating a ratio of 1:2. However, the duration of syllables that contained these cues were approximately 200 and 300 milliseconds. If infants were trying to discriminate syllables that differed minimally in duration, the ratio would have been 2:3, which exceeds the discrimination threshold for younger infants.

## 2.3 Input for duration-based phonemic contrasts

Another factor that could potentially contribute to the difficulty of duration-based phonemic discrimination is the availability of acoustic cues in the infants' input. As discussed above, durational variations in vowels and consonants occur in any language since duration of segments can vary with a number of factors, e.g. phrasal context (final lengthening, speech rate and so on). Yet, whether or not the durational differences are utilized phonemically differs from language to language. Accordingly, infants cannot know a priori that the durational differences are phonemic in the ambient language. This means that infants learning Japanese must notice not only that the second vowels of word pairs such as /toko/ (*bed*) versus /tokoo/ (*travel*) and /konpyuuta/ (computer) versus /konpyuutaa/ (computer) differ in the durations of the final vowels, but also that this difference corresponds to separate word meanings in the first pair but not in the second.

In order for infants to learn which contrasts are relevant in their language, at least two requirements should be met. The first requirement is that contrastive sounds should be acoustically different from each other (Werker et al. 2007). In the case of phonemic duration, long vowels should be significantly longer than short vowels. The second requirement is that the frequency distribution of the relevant acoustic cue should make the difference salient enough in the input. Infants should be able to learn that duration is contrastive in Japanese by determining the number of modes in the distribution of vowel durations without any knowledge about vowel identity (Vallabha et al. 2007; Maye and Gerken 2000; Maye, Werker and Gerken 2002). Bion et al. (2013) examined these two requirements of phonemic learning in the vowel durations in Japanese. Analyses were done in two steps. First, it was examined whether there are reliable differences in the duration of long and short vowels in Japanese infant-directed speech (IDS). Second, it was examined whether the frequency distribution of vowels from over 11 hours of naturalistic recordings provides enough evidence that duration is phonemic.

For analysis, Bion et al. (2013) used data from an IDS corpus, originally recorded for the RIKEN Japanese Mother-Infant Conversation Corpus (Mazuka, Igarashi and Nishikawa 2006). The speech samples were provided by 22 Japanese-speaking mothers from the Metropolitan Tokyo Area, while they interacted with their 18–24 month old infants. The same mothers were also recorded while they conversed with an adult experimenter in order to elicit adult-directed speech (ADS). The IDS recordings consisted of a total of 11 hours of speech, approximately 50,000 words.

Acoustic analyses were performed using Praat (Boersma 2002). For each vowel, we added information about phonemic length (short or long), height (high, mid, or low), and whether the vowel was immediately followed by a word boundary and/or intonational boundary. Adult-directed speech (ADS) fragments were removed from the IDS recordings. Singing, laughing, coughing, onomatopoeia and fragments the transcriber could not understand were also removed. Whether a vowel was long or short was decided on the basis of the lexical item intended by the speaker. If the transcriber hears the word /oka:san/ mother), the phonemic duration of each of the three vowels is unambiguous. The first vowel is phonemically short (i.e., it contains one mora), the second vowel is long (i.e., it contains two morae), and the third vowel is short. The transcriber would then label each vowel as phonemically short or long and mark its start and end on the audio file, from which duration could be subsequently computed. In the rare cases in which the lexical item was not easily recognizable, it was marked as such and not included in our analyses. In order to investigate whether there are reliable differences in the duration of short and long vowels in Japanese, the average duration of short and long vowels for each of the five Japanese vowels (i.e., a, e, i, o, u) were compared. Towards this end, mean values for the duration of short and long vowels were computed for each mother separately for each of the five oral vowels. As shown in Figure 4, it was found that long vowels

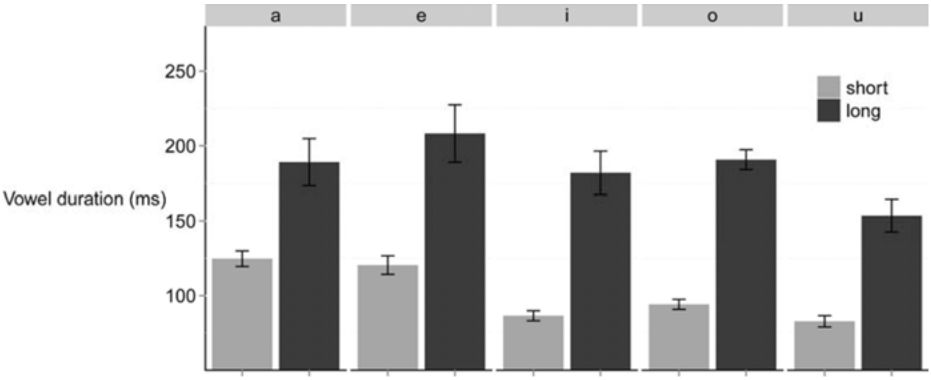



**Figure 4:** Mean duration of short and long vowels in Japanese IDS. The difference in duration between short and long vowels is reliable and the effect size is large. The error bars represent the standard error of the mean for each vowel across participants. (Bion et al. 2013: 3, Figure 1, The original color figure was modified to be legible in black and white.)

were reliably longer than short vowels independently of the vowel produced. When the average duration of long and short vowels were compared, long vowels were in fact significantly longer than short vowels. This was true for every one of 22 mothers, replicating the result of a previous study by Werker et al. (2007).

As discussed above, however, infants are not born with the knowledge of whether the durational differences are phonemic in their language. In the second analysis, we examined whether the durational differences of long versus short vowels are accessible in the input of infants. Figure 5 shows histograms with all the vowels in Bion et al.'s IDS sample. A visual inspection of this figure reveals two findings: most of the vowels in this corpus are short, and there is complete overlap in the distribution of short and long vowels in the input for each of the five oral vowels. In fact, 94% of the vowels in Japanese are short (27,561), and only 6% of them are long (1,942), a pattern replicated for each of the 22 speakers individually.

To test the generalizability of our findings, two additional ADS corpora were analyzed. The same 22 mothers from the IDS study were also recorded whilst speaking to an adult experimenter. Using the same coding scheme, it was found that 25,071 vowels (92.6%) were short while only 2,018 vowels were long. In addition, the vowels in the Corpus for Spoken Japanese (CSJ) (Maekawa 2003) were analyzed. Likewise, there were 449,494 short vowels (90.3%) and only 48,032 long vowels. These additional analyses confirm our finding that around 90% of vowel tokens from spontaneously spoken Japanese are short. Kunnari, Nakai and Vihman (2001) found a similar difference in base-rate between geminate and singleton consonants, with 93% of Japanese consonants consisting of a single mora (singleton). Similarly, the proportion of singleton consonants in the RIKEN IDS corpus was also 93%.

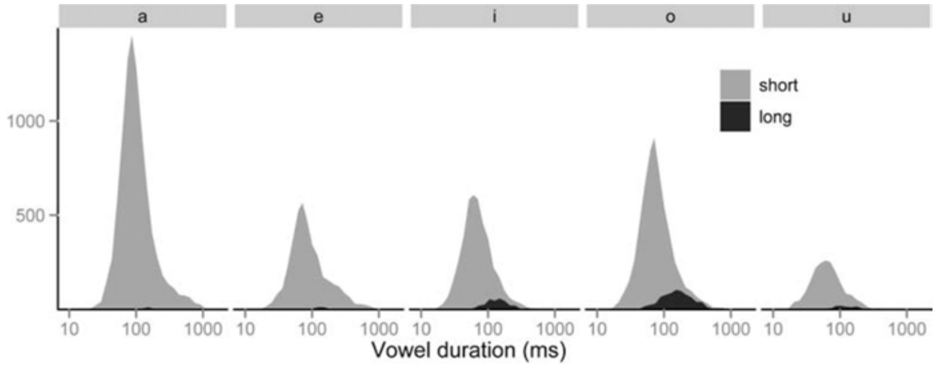



**Figure 5:** The stacked frequency distribution of vowel duration in Japanese IDS. Ninety-four percent of the vowels in our corpus are short, and there is a complete overlap in the distribution of short and long vowels. This kind of input is problematic for simple distributional learning models. (Bion et al. 2013: 3, Figure 2, The original color figure was modified to be legible in black and white.)

The above analyses were based on token frequencies; i.e., if the same word *mama* (mother) was repeated 6 times, the vowel /a/ was counted as occurring 12 times. To examine whether a different pattern would emerge in type frequency analysis, the frequencies of long and short vowels in different words were calculated. The results revealed the same pattern: 94% of all vowels were short while only 6% of the vowels were long.

A hierarchical regression analysis confirmed that other factors such as vowel height, phrasal positions, individual differences among mothers and individual words in which the vowels appeared also contributed to the vowel duration. The effects of these and several additional factors were examined to determine whether they would change the distribution of the long and short vowels, including normalizing vowel duration into z scores for each speaker, taking into account the duration of the following and preceding vowels, and computing separate distributions for vowels at intonational or word boundaries. However, these factors did not change the unimodal distribution due to the large differences in base-rate between short and long vowels.

Bion et al. (2013) make the crucial point that even if there are reliable differences in the durations of long and short vowels and infants may be capable of detecting such durational differences, the fact remains that the complete overlap of the durations of long and short vowels in the distribution of the input makes it impossible for Japanese infants to determine from the duration of the vowels *per se* that vowel duration is phonemic in Japanese. This implies that a simple distributional analysis of input (Vallabha et al. 2007; Maye, Werker and Gerken 2002) would not suffice as a mechanism of learning the duration-based vowel categories in Japanese.

# 3 Lexical-level prosody

Another property of Japanese phonology that presents a unique challenge for infants trying to acquire the language is lexical-pitch accent. Lexical-level prosody, which includes lexical stress in languages such as English and German, tones in languages such as Chinese, Vietnamese and Thai, and lexical-pitch accent in Japanese, uses acoustic cues that are typically used for prosody, such as pitch, duration and amplitude, to distinguish lexical meaning. Like the duration-based phonemic contrasts discussed in Section 2, the acquisition of lexical-level prosody requires infants to learn not only to distinguish the relevant acoustic cues, but also whether the specific differences are lexical or prosodic.

In Tokyo Japanese, a pair of two-syllable homophones are distinguished by a pitch-accent that either follows a high-to-low (HL) or low-to-high (LH) pitch pattern such as ha'shi (HL: 'chop stick') versus hashi' (LH: 'bridge'). This situation is similar to how tones are used in Chinese and Thai, in that changes in pitch contours mark lexical meaning. Unlike tones, however, Japanese lexical-pitch accent marks the position of an accented syllable within a word, while every syllable in a tone language is marked individually with a tone and the pitch changes occur within each syllable.

## 3.1 Discrimination of lexical-pitch accent

As in the case for duration-based phonemic contrast, the first question that needs to be addressed is how Japanese infants' ability to discriminate the lexical-pitch accent develops. It is possible that this ability develops in the enhancement pattern, similar to duration-based phonemic contrasts. The motivation for this comes from the fact that, like duration-based phonemic contrasts, lexical-level prosody involves the dual task of discriminating the relevant cues and distinguishing lexical and prosodic use of the same acoustic cues. Alternatively, it is possible that lexical-pitch accent is discriminable from early on. Support for this alternative comes from cross-linguistic studies. In a study with French neonates, Nazzi, Floccia and Bertoncini (1998) found that the infants could discriminate the Japanese lexical pitch-accent difference between HL and LH. Studies on English- and French-learning infants' discrimination of tones showed that they were able to discriminate the foreign (Thai) tones at 6 months of age but lost this ability by 9 months of age (Mattock and Burnham 2006; Mattock, Molnar, Polka and Burnham 2008). More recently, Yeung, Chen and Werker (2013) tested Mandarin, Cantonese, and English learning infants on their discrimination of Cantonese and Mandarin tones and reported the maintenance/decline pattern of development; younger infants were able to discriminate both native and non-native tones, while older infants were able to discriminate only the native contrasts.

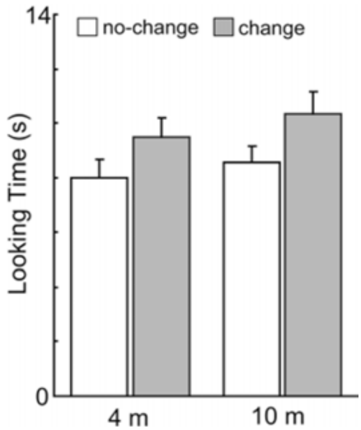



**Figure 6:** Average (mean) looking times (with standard error bars) during the no-change and change trials for each age group (Sato, Sogabe and Mazuka 2010b: 2506, Figure 1). No-change trials in this figure are the same as the same trials, and change trials as the switch trials in Figures 1–3.

To test these possibilities, Sato, Sogabe and Mazuka (2010b) tested 4- and 10-month old Japanese infants on their discrimination of lexical pitch-accent changes (HL vs. LH) embedded in Japanese disyllabic words. As in the experiments reported above in Section 2, the modified visual habituation method was used. The stimuli consisted of 14 existing disyllabic Japanese word pairs that minimally differ in pitch accent, such as *ame*' ('candy') – *a'me* ('rain'), *kiri*' ('fog') – *ki'ri* ('paulownia tree') and *kame*' ('jag') – *ka'me* ('turtle'). These stimuli were adopted from Nazzi et al. (1998).

The results of the behavioral experiments are shown in Figure 6. Both 4- and 10-month-old Japanese infants were able to discriminate the word pairs that minimally differed in the lexical-pitch accent of HL vs. LH. This is consistent with the cross-linguistic studies on the discrimination of lexical-level prosody from French, English, and Chinese learning infants, and shows that Japanese infants start out with an early sensitivity to the lexical pitch-accent contrasts. Interestingly however, this is a different pattern from the duration-based phonemic contrasts discussed above in Section 2.

## 3.2 Infants' brain activation for lexical-pitch accent

Behaviorally, Japanese 4- and 10-month-olds alike discriminated the lexical pitch-accent contrasts, and on the basis of their behavioral responses, both groups appeared to process the lexical pitch-accent contrasts in the same way. Yet, as discussed in Section 1, infants' ability to perceive speech sounds goes through significant changes during the first year of life. It has been argued that the observed changes in infants' behavior reflect the way infants process speech stimuli. That is,

whereas infants initially process speech stimuli via general auditory processing, by the second half of the first year they begin to process speech in their own language in a specific way (Kuhl 2004; Werker and Tees 1992, 2005). This process is sometimes called "reorganization" (Werker and Tees 1992). The proposal, thus far, has been built on the basis of cumulative cross-linguistic studies showing that infants begin to lose discrimination ability to foreign contrasts during the second half of the first year. The implication drawn from the behavioral data alone is indirect, however, since the loss of discrimination for foreign contrasts does not *a priori* entail that infants are processing native contrasts as "linguistically relevant".

Data from infants' brain activation patterns may help shed light on this process. It is thought that reorganization is likely to be associated with underlying neural development for processing speech (Kuhl 2004; Werker and Tees 2005). According to event-related potential (ERP) studies, younger infants show similar mismatch negativity (MMN) patterns in response to both native and non-native phonemic contrasts, whereas older infants show either smaller or no MMN responses to non-native contrasts (Cheour et al. 1998; Rivera-Gaxiola, Silva-Pereyra, and Kuhl 2005; see also Sakamoto's chapter in this volume on ERP). Although the MMN results provide evidence that younger and older infants are processing the native and non-native contrasts differently, they do not necessarily indicate that the older infants are distinguishing the native contrast as "linguistically relevant".

Evidence for this is likely to come from imaging studies that test functional lateralization of speech processing. It is well known that in adults the left and right cerebral hemispheres work differently for speech processing: the left hemisphere is more heavily involved in processing segmental contrasts in one's native language, and the right hemisphere typically processes prosodic cues including affective prosody. Bilateral activation is seen in the processing of non-speech or non-native contrasts (Buchanan et al. 2000; Jacquemot et al. 2003; Näätänen et al. 1997; Ross 1981; Schirmer and Kotz 2006; Tervaniemi et al. 1999; van Lancker 1980; Vouloumanos et al. 2001; Zatorre et al. 1992). If, as the reorganization hypothesis predicts, infants begin to process linguistically relevant speech stimuli differently from other auditory stimuli post-reorganization, we may observe a shift in hemispheric dominance between the younger and older infants as they learn during the first year of life that a particular contrast is linguistically relevant in their language.

Lexical level prosody is ideally suited to test this prediction. As discussed above, lexical level prosody, such as lexical pitch-accent in Japanese and tones in Chinese and Thai, uses prosodic acoustic cues (e.g., pitch changes) to distinguish lexical meaning. Brain activation for these stimuli seems to be functionally determined; when the pitch cues for the lexical prosody are processed as linguistically relevant, left-lateralized activations are found, and bilateral or no left dominance activation is seen when the same cue is processed non-linguistically (Gandour et al. 2000; Gandour, Wong and Hutchins 1998; Klein et al. 2001; Sato, Sogabe, and Mazuka 2007; Wang et al. 2003). The use of lexical pitch-accent stimuli allowed us to test

linguistic and non-linguistic contrasts within one language by presenting the same pitch cue in word pairs and pure tones.

It is not the case that the two hemispheres of young infants are symmetrical for processing auditory or speech stimuli. A number of studies have reported a differential involvement of the two hemispheres for speech stimuli from early in infancy. In neonates and 3-month-olds, a stronger activation in the left hemisphere has been reported for regular speech than for backward speech or silence (Dehaene-Lambertz, Dehaene and Hertz-Pannier 2002; Peña et al. 2003), and stronger right-side activation was reported in 3-month-olds when regular speech was compared to speech with flattened intonation (Homae et al. 2006). Some ERP studies have also shown the presence of an early (2 to 4 months) left dominance for some segmental processing, such as /ba/ versus /ga/ (Dehaene-Lambertz 2000; Dehaene-Lambertz and Baillet 1998; Dehaene-Lambertz and Dehaene 1994). Yet, the differences and similarities between this early asymmetry and full-fledged functional lateralization of language processing in adults are not yet well understood.

If the behavioral differences accompanying reorganization in infants are reflected in the functional lateralization for processing speech stimuli, we predict that in discriminating linguistically relevant (native phonemic) and non-relevant (non-native) contrasts, younger infants should process both in the same way (i.e., no left-hemisphere dominance expected), whereas older infants should show a left-side dominance only for linguistically relevant contrasts. To test this prediction, Sato, Sogabe and Mazuka (2010b) examined the brain activation of 4- and 10-month-old Japanese infants while they listened to the lexical pitch-accent stimuli. Hemodynamic responses were recorded with a multi-channel Near Infrared Spectroscopy (NIRS) system (ETG-4000, Hitachi Medical Co., Japan), which uses near infrared lasers at two wavelengths (695 and 830 nanometer). NIRS non-invasively measures relative changes in the concentration of hemoglobin (Hb) in localized brain tissues without the loud noises that are associated with MRI. It requires minimum constraints on participants, and is therefore well-suited for studying infants (Homae et al. 2006; Minagawa-Kawai et al. 2007; Peña et al. 2003). We focused on the Oxygenated-Hb (Oxy-Hb) responses, which represent cerebral blood oxygenation.

There were two sets of stimuli: (i) word stimuli that were used in the behavioral experiment discussed in Section 3.1 and in the previous study (Sato, Sogabe and Mazuka 2007), and (ii) 14 pure-tone pairs (HL vs. LH) that were created by extracting the fundamental frequencies from word tokens used in the lexical pairs (Sato, Sogabe and Mazuka 2007).

Infants were tested in two conditions in a block design. In the word condition, the baseline block (20 seconds or 25 seconds) contained a sequence of either HL or LH words repeated approximately every 1.25 s. During the test block (10 seconds), participants were presented with words featuring both pitch patterns. The HL and LH pattern words were presented in a pseudo-random order with equal probability.

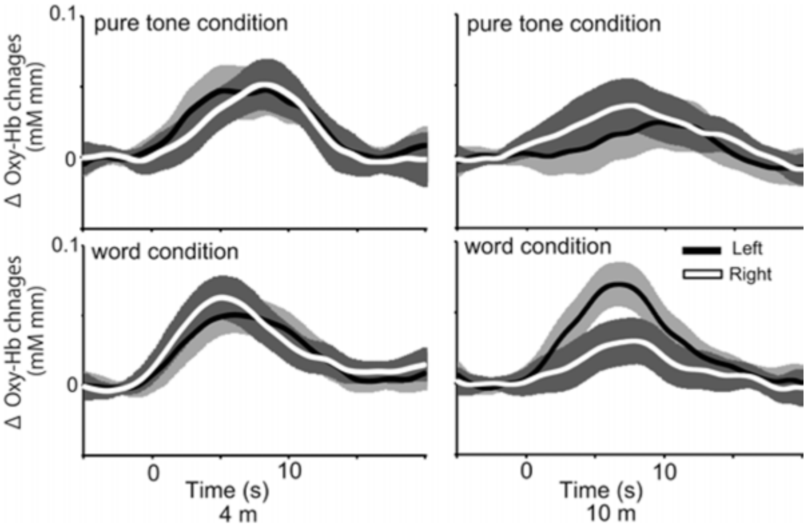



**Figure 7:** The grand averages for time courses of the Oxy-Hb changes from one channel of the maximum response in the left (black lines) and the right (white lines) hemispheres under two conditions in each age group. Light and dark gray areas above and below the traces indicate standard errors. The marks at zero and 10 on the horizontal lines show the beginning and end of a test block, respectively. (Sato, Sogabe and Mazuka 2010b: 2508, Figure 4, The original color figure was modified to be legible in black and white.)

The pure-tone (PT) condition was similar to the word condition, except for the presentation of the pure-tone stimuli.

Figure 7 shows the mean time course of Oxy-Hb concentration changes of both age groups for left and right ROI channels in the two conditions. The y-axes are the grand averages of Oxy-Hb responses from all participants in each condition. The lateralization pattern under the word condition differed between the two age groups. Whereas 4-month-olds showed similar Oxy-Hb responses under both conditions, the responses of 10-month-olds differed between the conditions. The overall activation levels between pure tone and word condition were not significantly different from each other in either age group. This showed that the Oxy-Hb changes elicited by pure tone stimuli were on average comparable to that of word stimuli for these infants. The overall activation level did not differ between left and right side either, suggesting that the stimuli in the present experiment elicited comparable levels of Oxy-Hb changes in either side of the brain.

The behavioral experiments demonstrated that 4- and 10-month-old infants did not differ in their behavioral responses to HL versus LH stimuli. The neural responses to the HL versus LH stimuli, however, revealed an important difference between the two groups. Although both age groups showed higher activation levels in the test

blocks than in the baseline blocks in both the pure tone condition and the word condition, indicating they detected the pitch changes in both types of stimuli, their responses to two types of stimuli in the two hemispheres differed between the two age groups: 10-month-old infants showed stronger left-hemisphere hemodynamic responses to the pitch changes embedded in word forms, but no left-side dominance in responses to pure tone stimuli, whereas 4-month-olds showed bilateral responses to both types of stimuli. Older infants' responses showed similar patterns to those seen in Japanese adults in a previous study (Sato, Sogabe and Mazuka 2007). The behavioral experiment confirmed that any changes we observed between 4 and 10 months of age were not due to changes in the ability to discriminate the stimuli. The change was found only in how the two brain hemispheres processed these contrasts.

As discussed above, it has been proposed that how infants process speech stimuli goes through a qualitative shift, from general auditory processing to specifically attuned processing for linguistically relevant contrasts (Kuhl 2004; Werker and Tees 1992, 1999, 2005). Sato, Sogabe and Mazuka (2010b) provided evidence to show that this reorganization is linked to the functional lateralization of speech processing; older infants process speech contrasts as "linguistically relevant," indicated by the (adult-like) left-hemisphere dominance in brain activation, while younger infants did not show this left-side dominance, indicating that the contrasts are yet to become linguistically relevant at that age.

The results of Sato, Sogabe and Mazuka (2010b) provide further evidence that research on Japanese phonological development can contribute significantly to address important disputes in the field. Since the early studies by Broca (1861) and Wernicke (1874), it has been well documented that the left hemisphere plays a more dominant role than the right in processing language and speech stimuli. Still, there has been an active debate as to what drives this asymmetry. The dominant view in the field has been that it is the linguistic function of the stimuli that drives the lateralization; we will call this a functional account. Important evidence for this account comes from the processing of lexical tones in Chinese and Thai (Klein et al. 2001; Gandour, Wong and Hutchins 1998; Gandour et al. 2000, 2002, 2003; Wong et al. 2004). In a positron emission tomography (PET) study, Gandour et al. (2000) found that Thai adults show left-hemisphere dominance to Thai tones, while such asymmetry is not found in Chinese or English speakers, demonstrating that the linguistic function of the tone is sufficient to drive a left-hemisphere advantage. On the other hand, it has been proposed that the observed asymmetry is not a lateralization driven by the function of the speech signal but rather reflects the physical properties of speech signals; auditory stimuli with slow acoustic transitions, such as pitch change, are preferentially processed in the right hemisphere whereas rapidly changing sounds, like consonants, are preferentially processed in the left (Zatorre and Belin 2001; Zatorre, Belin and Penhune 2002). We will call this an

acoustic account. A review of the literature suggests that both of these factors may contribute to the functional lateralization of language and speech stimuli in the adult brain. But as Zatorre and Gandour (2008) pointed out, much is still to be discovered about how these factors interact.

Developmental research could provide important insight to this debate. As discussed above, cumulative data from electrophysiological and imaging studies with young infants suggests that the two hemispheres are not symmetrical in processing speech and other auditory stimuli from early in infancy (Dehaene-Lambertz 2000; Dehaene-Lambertz and Baillet 1998; Dehaene-Lambertz and Dehaene 1994, Dehaene-Lambertz, Dehaene and Hertz-Pannier 2002; Dehaene-Lambertz and Gliga 2004; Homae et al. 2006; Peña et al. 2003). Infants at this young age are not likely to have learned much about the specific characteristics of their native language segments, while they may have already learned some prosodic properties. Thus, these data show that at least some asymmetry is already present prior to "reorganization." Still, we cannot determine whether the observed asymmetry is derived purely from the physical properties of the stimuli (e.g., fast or slow transition) or is functionally linked to the fact that some of the stimuli were human speech. Moreover, as these studies have focused on very young infants, we do not know whether the said asymmetry changes as infants learn the linguistic relevance of particular speech stimuli.

In Sato, Sogabe and Mazuka (2010b), a change was observed in the left-hemisphere dominance for pitch-accent embedded in word forms between 4 and 10 months of age. Compared to the early asymmetry data discussed above, the emergence of the left-side advantage in Sato, Sogabe and Mazuka (2010b) is relatively late. This allows dissociation of the acoustic and developmental factors that could contribute to the emergence of a left-hemisphere advantage. The phonemic contrast used for this study was the lexical pitch-accent of Japanese. Similar to the Gandour et al. (2000) study with tones, the acoustic cue for the lexical pitch-accent in Japanese is a slow transition of the pitch. This type of cue is typically associated with bilateral activation or a right-hemisphere advantage. Had a typical phonemic contrast with consonants or vowels involving fast formant transitions been used, a left-hemisphere advantage could have arisen on the basis of the inherent acoustic properties of the stimuli independent of its linguistic function (Zatorre and Belin 2001; Zatorre Belin and Penhune 2002). It was found that 4-month-old infants did not have a left-hemisphere advantage for either the pure tone stimuli or word form stimuli, showing that the LH advantage found in 10-month-old infants is not attributable to the acoustic property of the pitch-accent stimuli. Instead, it should be due to the developmental changes that occurred to infants between 4 and 10 months of age.

# 4 Phonological grammar

Languages differ not only in the inventory of particular segments (consonants and vowels), but also in the phonological grammar which governs how segments can be assembled in a linear sequence. For instance, in languages that allow only a restricted set of syllable types, like Japanese, most consonants (C) are obligatorily followed by a vowel (V). In others with more complex syllabic types, such as French or English, one can find long series of consonants (e.g. CCCVCC as in “strict”). Several studies have shown that adults have trouble perceiving illegal sequences of segments, and even tend to misperceive segments in order to “repair” these sequences (Hallé et al. 1998). For instance, unlike French adults, Japanese adults perceive a nonword like “abna” as “abuna”, inserting the illusory (epenthetic) vowel /u/ to break up the illegal consonant cluster (Dupoux et al. 1999; Dupoux et al. 2001; Dehaene-Lambertz, Dupoux and Gout 2000; Jacquemot et al. 2003; Kubozono 1995; Vance, 1987). As a part of learning the sound system of a language, infants must learn the phonological grammar of the language along with segmental and prosodic properties. As a consequence of acquiring the phonological grammar of Japanese, infants are predicted to begin hearing the phonologically-induced illusions as Japanese adults do. Yet, very few studies have examined the acquisition of phonological grammar.

On the one hand, it is possible that phonological grammar is acquired along the same time course as segmental categories, as discussed above. Alternatively, however, it could take a different developmental trajectory. This is because the acquisition of a phonological grammar is potentially more complex than the acquisition of segmental categories. Infants must detect and remember sequences of segments and the contexts in which they occur. Indeed, most formal models of phonological acquisition assume that learning takes place through the comparison of underlying lexical representations and surface word forms (Gildea and Jurafsky 1996; Tesar and Smolensky 1998, 2000). This presupposes that children have to have acquired a lexicon beforehand, which contains not only word forms but also underlying forms. Supposedly, in order to recover underlying forms, children would need to have access to at least some morphological alternations, which would imply a rather complex lexical knowledge that potentially includes word meanings. This in turn predicts that phonological illusions should emerge only after children have acquired a large enough lexicon to enable robust induction of the grammar. In contrast, other models propose that incomplete but robust fragments of the native phonology can be bootstrapped from a bottom-up analysis of distribution of segments (Peperkamp and Dupoux 2002; Peperkamp 2003). These models predict that such illusions might appear as soon as infants acquire their native inventories of segments, at around one year of age.

Experimental studies showing that young infants pay attention to sequential distributional regularities provide supporting evidence for the latter model, e.g., 9-month-olds prefer to listen to words containing legal or frequent sequences of segments, rather than illegal or infrequent ones (Jusczyk and Luce 1994; Jusczyk et al. 1993), and by 16.5 months of age, they are able to learn phonotactic regularities from a brief exposure (Chambers, Onishi and Fisher 2003). Yet, showing that infants have become sensitive to the phonotactic regularities of their language is not sufficient to determine whether the phonologically-induced illusion is also experienced. Indeed, previous studies on infants' phonotactic acquisition have shown that infants are able to discriminate between the prototypical versus unusual sequences, since they prefer to listen to prototypical sequences (Jusczyk et al. 1993; Jusczyk and Luce 1994). But it is a quite different phenomenon to "repair" a non-legal sequence in order to make it congruent with the native language. Although different languages use different repair strategies (LaCharité and Paradis 2005), the phonologically-induced /u/ illusion in Japanese is a perfect case to study this, since it has been extensively studied in adults (Dupoux et. al. 1999; Dupoux et.al. 2001; Dehaene-Lambertz, Dupoux and Gout 2000; Jacquemot et al. 2003).

To shed light on the *mechanisms* that may lead to the acquisition of such a phonological illusion in Japanese speakers, Mazuka et al. (2011) conducted a cross-linguistic study comparing Japanese and French learning infants. In 3 experiments, French and Japanese infants were tested at 8 and 14 months of age using the modified visual habituation paradigm (the same method used in the behavioral experiments discussed in Sections 2 and 3). In the first two experiments, the stimuli were 8 pairs of nonsense words in the form of $V_1C_1C_2V_2$, such as /abna/, /ebzo/, /obda/, and their counterparts in $V_1C_1uC_2V_2$, such as /abuna/, /ebuzo/, /obuda/. Each pair differs minimally in the presence of /u/ between the two consonants. The results revealed that although 8-month old infants in both languages discriminated these pairs reliably, only French infants discriminated them at 14-months. Japanese infants were at chance level in discriminating these pairs by this age, as shown in Figure 8. As a control, 8- and 14-month-old Japanese infants were also tested in their discrimination of a single word pair; /abna/ versus /abuna/. In this case, both 8- and 14-month-old Japanese infants reliably discriminated the pair.

The results of these experiments revealed a robust cross-linguistic difference in the discrimination of VCCV versus VCuCV stimuli by French and Japanese 14-month-old infants. The performance of Japanese infants is most consistent with the account that the phonologically-induced /u/ illusion (perceptual epenthesis) is already in place by the age of 14 months. The fact that Japanese infants were able to discriminate /abuna/ versus /abna/ in the absence of high phonetic variability rules out the possibility that the failure of 14-month-olds in the high phonetic variability condition was the result of their inability to exploit the acoustic cues that distinguish the two types of stimuli, or to perform the experimental task in the presence of illegal stimuli.

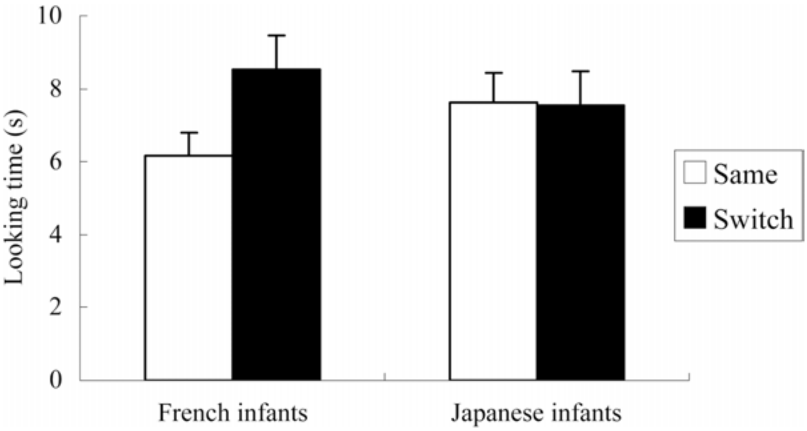



**Figure 8:** Average looking times during Same and Switch trials for French and Japanese 14 month-old infants in the high phonetic variability condition. Error bars represent the standard error of the difference of Switch and Same trials. (Mazuka et al. 2011: 696, Figure 2)

Having managed to learn the prototypical phonotactic patterns of one's mother tongue *on its own* does not explain the failure to discriminate the illegal clusters from prototypical sequences. Previous studies on the acquisition of phonotactic knowledge in infants relied on the head-turn preference paradigm and found that by 9 months, English-learning infants show a preference for the phonotactic patterns that are legal or more frequent in their language over illegal or less frequent ones (Jusczyk et al. 1993; Jusczyk and Luce 1994). These studies demonstrate that infants are able to discriminate the legal (or more frequent) sequences from illegal (or less frequent) sequences, even though they have already learned which ones are more prototypical in their language. The results of Mazuka et al. (2011) thus go beyond the mere sensitization to language-specific phonotactic patterns.

Instead, their results can be most straightforwardly accounted for if we assume that Japanese infants have come to experience the phonologically-induced /u/ illusion by 14 months of age. Japanese adults insert an epenthetic vowel /u/ when they attempt to produce foreign words that contain consonant clusters that are illegal in Japanese. When they hear /abna/, they report "hearing" an illusory vowel /u/, and as a result, they find it difficult to discriminate /abna/ from /abuna/. French adults, in contrast, easily produce /abna/ and /abuna/ distinctively, they do not "hear" /u/ in /abna/, nor do they have difficulty discriminating the pair (Dupoux et al. 1999). While the Japanese and French infants did not differ significantly from each other at 8 months, they behaved significantly differently at 14 months. French 14-month-olds, like French adults, showed no difficulty discriminating the pair, while Japanese 14-month-olds, like Japanese adults, were unable to discriminate the pair.

In sum, Mazuka et al. (2011) found that the acquisition of perceptual epenthesis takes place by 14 months of age, which is on a par with that of segmental categories (Werker and Tees 1984; Polka and Werker 1994; Kuhl et al. 1992; Stager and Werker 1997) and of the statistical regularities of segment sequences (Chambers, Onishi, and Fisher 2003; Kajikawa et al. 2006; Jusczyk et al. 1993; Jusczyk and Luce 1994). Acquisition presumably takes place primarily using the statistical learning mechanisms which only refer to distributions of sounds within utterances, before the information from a large lexicon becomes available (Maye, Werker and Gerken 2002; Chambers, Onishi and Fisher 2003; Peperkamp et al. 2006; Anderson, Morgan and White 2003; White et al. 2008). Still, vowel epenthesis has traditionally been analyzed as an integral part of phonological grammar (Rose and Demuth 2006; Uffmann 2006). Indeed, the repair of illegal syllabic structures is not universal but depends on language-specific properties; depending on the language, repairs are done through deletions, epenthesis or segmental change (LaCharité and Paradis 2005). But if epenthesis depends on acquisition of the full phonological system, it should be observed only after the acquisition of a sufficient number of words (Gildea and Jurafsky 1996; Tesar and Smolensky 1998, 2000). The fact that it arises by 14 months makes this claim unlikely and reinforces the view that infants first undergo a statistical learning phase, where incomplete but robust fragments of the phonological grammar are inferred on the basis of a distributional bottom-up analysis of continuous speech, which may or may not be segmented into word-sized units (Peperkamp and Dupoux 2002; Peperkamp 2003; Peperkamp et al 2006). Such an early acquisition would provide the foundations for a more complete lexical-based learning (Pierrehumbert 2003). It would also result in long-lasting perceptual tuning and thereby explain the difficulties adults have in perceiving foreign languages.

# 5 Conclusion

Within a short time after birth, infants become attuned to the phonological system of the ambient language. This allows them to begin the next step of language acquisition, which involves learning to comprehend and produce words, sentences and beyond. Compared to adults, who struggle to learn the sound system of a foreign language, the speed and efficiency at which infants acquire the sound system of their language has attracted extensive research into how such a feat is made possible. Our own research attempts to approach this question by focusing on how Japanese infants learn some specific properties of Japanese phonology, which differ from those of English and other European languages that have dominated research in this field. The present chapter discussed three lines of research in this approach. The first set of studies involved duration-based phonemic contrasts, such as the distinctions between long and short vowels and between singleton and geminate

obstruent contrasts. We have demonstrated that unlike the majority of phonemic contrasts that have been studied previously, discrimination of duration-based phonemic contrasts is difficult for young infants, who take until well into the second half of their first year before becoming able to make the distinction. The analyses of infant-directed speech revealed that the nature of input may be one factor that contributes to this difficulty.

The second study examined the acquisition of lexical pitch-accent. Behaviorally, infants were able to discriminate the lexical pitch-accent from an early age, which is different from duration-based phonemic contrasts. When brain activation to the lexical-pitch accent stimuli was examined in a NIRS study however, it was found that 4- and 10-month-old infants showed a different pattern in their left and right-hemispheres. This suggested that while infants' brains start out processing the phonological segments of their native language in the same way as non-native or non-linguistic stimuli, they begin to process those in their own language as linguistically relevant by 10 months of age. Duration and pitch are two of the fundamental acoustic properties of speech, and infants in any language must learn how such cues are utilized in their language. By taking advantage of duration-based phonemic contrasts and the lexical pitch-accent of Japanese, which also uses these cues lexically, we were able to reveal that i) infants acquire quality- and quantity-based phonemic contrasts differently, ii) although infants are initially sensitive to pitch changes embedded in lexical pitch-accent, the same cannot be said for duration changes embedded in vowels and consonants, and iii) functional lateralization in infants' brain activation shows that infants begin processing phonemic contrasts in their native language as linguistically relevant by the second half of their first year.

The third study dealt with Japanese infants' perception of the phonologically induced illusionary vowel /u/. It revealed that by 14 months or so, Japanese infants are adult-like in hearing epenthetical vowels in words like /abna/. This suggests that the tuning of one's ear to the sound system of the native language by 14 months is robust enough to force the language-learner to hear a vowel that is not there. It may explain why it is so difficult for adults to learn the sound system of a foreign language.

## Acknowledgments

The studies discussed in this chapter were originally published in the papers referenced, and they should be referred to for further details of methodology and results. The research reported in this paper was in part supported by Grants-In-Aid for Scientific Research (Kakenhi, the Japanese Ministry of Education, Culture, Sports, Science and Technology No. 21610028 and No. 232420) We thank Robert Grayson and Risako Omori for their help for the preparation of the manuscript.

# References


Anderson, Jennifer L., James L. Morgan and Katherine S. White. 2003. A statistical basis for speech sound discrimination. *Language and Speech* 46. 155–182.

Beckman, Mary E. and Jan Edwards. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In John Kingston and Mary. E. Beckman (eds.), *Papers in laboratory phonology 1: Between the grammar and physics of speech*, 152–178. Cambridge: Cambridge University Press.

Best, Catherine T. and Gerald W. McRoberts. 2003. Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech* 46(2–3). 183–216.

Best, Catherine T., Gerald W. McRoberts and Nomathemba M. Sithole. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of experimental psychology. Human perception and performance* 14(3). 345–360.

Bion, Ricardo A. H., Koki Miyazawa, Hideaki Kikuchi and Reiko Mazuka. 2013. Learning phonemic vowel length from naturalistic recordings of Japanese infant-directed speech. *PLOS ONE* 8(2). e51594.

Boersma, Paul. 2002. Praat, a system for doing phonetics by computer. *Glot International* 5. 341–345.

Brannon, Elizabeth M., Sumarga Suanda and Klaus Libertus. 2007. Temporal discrimination increase in precision over development and parallels the development of numerosity discrimination. *Developmental Science* 10(6). 770–777.

Broca, Paul. 1861. Remarques sur le siège de la faculté du la language articulè, suivies d'une observation d'aphemie (perte de la parole). *Bulletin de la Société Anatomique (Paris)* 36. 330–357. (Translation: Jules Kann (1950). A translation of Broca's original article on the location of the speech center. *Journal of Speech and Hearing Disordors* 15. 16–20.)

Buchanan, Tony W., Kai Lutz, Shahram Mirzazade, Karsten Specht, N. Jon Shah, Karl Zilles and Lutz Jäncke. 2000. Recognition of emotional prosody and verbal components of spoken language: An fMRI study. *Cognitive Brain Research* 9(3). 227–238.

Chambers, Kyle E., Kristine H. Onishi and Cynthia Fisher. 2003. Infants learn phonotactic regularities from brief auditory experience. *Cognition* 87(1). B69–B77.

Cheour, Marie, Rita Ceponiene, Anne Lehtokoski, Aavo Luuk, Jüri Allik, Kimmo Alho and Risto Näätänen. 1998. Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience* 1(5). 351–353.

Cooper, William E. and Jeanne Paccia-Cooper. 1980. *Syntax and speech.* Cambridge, MA: Harvard University Press.

Dehaene-Lambertz, Ghislaine. 2000. Cerebral specialization for speech and nonspeech stimuli in infants. *Journal of Cognitive Neuroscience* 12. 449–460.

Dehaene-Lambertz, Ghislaine, Emmanuel Dupoux and Ariel Gout. 2000. Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience* 12(4). 635–647.

Dehaene-Lambertz, Ghislaine and Stanialas Dehaene. 1994. Speech and cerebral correlates of syllable discrimination in infants. *Nature* 370. 292–295.

Dehaene-Lambertz, Ghislaine, Stanislas Dehaene and Lucie Hertz-Pannier. 2002. Functional neuroimaging of speech perception in infants. *Science* 298(5600). 2013–2015.

Dehaene-Lambertz, Ghislaine and Sylvain Baillet. 1998. A phonological representation in the infant brain. *NeuroReport* 9(8). 1885–1888.

Dehaene-Lambertz, Ghislaine and Teodora Gliga. 2004. Common neural basis for phoneme processing in infants and adults. *Journal of Cognitive Neuroscience* 16. 1375–1387.

Dupoux, Emmanuel, Christophe Pallier, Kazuhiko Kakehi and Jacques Mehler. 2001. New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes* 5(16). 491–505.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier and Kacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25(6). 1568–1578.

Eek, Arvo. 1984–5. Problems of quantity in Estonian word prosody. *Estonian Papers in Phonetics*. 13–66.

Eilers, Rebecca E., Wesley R. Wilson and John M. Moore. 1979. Speech discrimination in the language-innocent and the language-wise: A study in the perception of voice onset time. *Journal of Child Language* 6(1). 1–18.

Eimas, Peter D., Einar R. Siqueland, Peter Jusczyk and James Vigorito. 1971. Speech perception in infants. *Science* 171(3968). 303–306.

Engstand, Olle. 1987. Preaspiration and the voicing contrast in Lule Sami. *Phonetica* 44(2). 103–116.

Fry, Dennis Butler. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27(4). 765–769.

Gandour, Jack, Donald Wong and Gary Hutchins. 1998. Pitch processing in the human brain is influenced by language experience. *NeuroReport* 9(9). 2115–2119.

Gandour, Jack, Donald Wong, Li Hsieh, Bret Weinzapfel, Diana Van Lancker and Gary D. Hutchins. 2000. A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience* 12(1). 207–222.

Gandour, Jack, Donald Wong, Mark Lowe, Mario Dzemidzic, Nalarin Satthamnuwong, Yunxia Tong and Xiaojian Li. 2002. A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. *Journal of Cognitive Neuroscience* 14(7). 1076–1087.

Gandour, Jack, Mario Dzemidzic, Wong Donald, Mark Lowe, Yunnxia Tong, Li Hsieh, Nakarin Satthamnuwong and Joseph Lurito. 2003. Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language* 84(3). 318–336.

Gildea, Daniel and Daniel Jurafsky. 1996. Learning bias and phonological rule induction. *Computational Linguistics* 22(4). 497–530.

Hallé, Pierre A., Juan Segui, Uli Frauenfelder and Christine Meunier. 1998. Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24(2). 592–608.

Hankamer, Jorge, Aditi Lahiri and Jacques Koreman. 1989. Perception of consonant length: Voiceless stops in Turkish and Bengal. *Journal of Phonetics* 17. 283–298.

Hansen, Benjamin B. 2004. Production of Persian geminate stops: Effects of varying speaking rate. In Augustine Agwuele, Willis Warren and Sang-Hoon Park (eds.), *Proceedings of the 2003 Texas Linguistics Society Conference*, 86–95, Somerville, MA: Cascadilla Proceedings Project.

Hirata, Yukari. 2004. Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics* 32(4). 565–589.

Hoequist, Charles. 1983. Syllable duration in stress-, syllable-, and mora-timed language. *Phonetica* 40. 185–203.

Homae, Fumitaka, Hama Watanabe, Tamami Nakano, Kayo Asakawa and Gentaro Taga. 2006. The right hemisphere of sleeping infant perceives sentential prosody. *Neuroscience Research* 54(4). 276–280.

House, Arthur S. and Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25(1). 105–113.

Jacquemot, Charlotte, Christophe Pallier, Denis LeBihan, Stanislas Dehaene and Emmanuel Dupoux. 2003. Phonological grammar shapes the auditory cortex: A functional magnetic resonance imaging study. *Journal of Neuroscience* 23(29). 9541–9546.

Jusczyk, Peter W., Angela D. Friederici, Jeanine M. I. Wessels, Vigdis Y. Svenkerud and Ann Marie Jusczyk. 1993. Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language* 32(3). 402–420.

Jusczyk, Peter W. and Paul A. Luce. 1994. Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33(5). 630–645.

Kajikawa, Sachiyo, Laurel Fais, Ryoko Mugitani, Janet F. Werker and Shigeaki Amano. 2006. Cross-language sensitivity to phonotactic patterns in infants. *Journal of Acoustic Society of America* 120(4). 2278–2284.

Klein, Denise, Robert J. Zatorre, Brenda Milner and Viviane Zhao. 2001. A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *Neuroimage* 13(4). 646–653.

Kubozono, Haruo. 1995. *Go-keisei to on'in-kōzō* [Word formation and phonological structure]. Tokyo: Kurosio Publishers.

Kuhl, Patricia K. 2004. Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience* 5(11). 831–843.

Kuhl, Patricia K., Erica Stevens, Akiko Hayashi, Toshisada Deguchi, Shigeru Kiritani and Paul Iverson. 2006. Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science* 9(2). F13–F21.

Kuhl, Patricia K., Feng-Ming Tsao, Huei-Mei Liu, Yang Zhang and Bart De Boer. 2001. Language/Culture/Mind/Brain. Progress at the margins between disciplines. *Annals of the New York Academy of Sciences* 935. 136–174.

Kuhl, Patricia K., Karen A.Williams, Franciso Lacerda, Kenneth N. Stevens and Bjorn Lindblom. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255(5044). 606–608.

Kunnari, Sari, Satsuki Nakai and Marilyn M. Vihman. 2001. Crosslinguistic evidence for acquisition of geminates. *Psychology of Language and Communication* 5(2). 13–24.

Kushnerenko, Elena, Rita Ceponiene, Vineta Fellman, Minna Huotilainen and István Winkler. 2001. Event-related potential correlates of sound duration: similar pattern from birth to adulthood. *NeuroReport* 12(17). 3777–3781.

LaCharité, Darlene and Carole Paradis. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36(2). 223–258.

Ladefoged, Peter. 1975. *A course in phonetics*. New York: Harcourt Brace Jovanovich College Publishers.

Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford, UK: Blackwell Publishers Ltd.

Lahiri, Aditi and Jorge Hankamer. 1988. The timing of geminate consonants. *Journal of Phonetics* 16. 327–338.

Lasky, Robert E., Ann Syrdal-Lasky and Robert E. Klein. 1975. VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology* 20. 215–225.

Lehiste, Ilse. 1966. *Consonant quantity and phonological units in Estonian*. Bloomington, IN: Indiana University dissertation.

Leppänen, Paavo H. T., Elina Pihko, Kenneth M. Eklund and Heikki Lyytinen. 1999. Cortical responses of infants with and without a genetic risk for dyslexia: II. Group effects. *NeuroReport* 10(5). 969–973.

Maekawa, Kikuo. 2003. Corpus of spontaneous Japanese: Its design and evaluation. *Proceedings of ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition Tokyo*. 7–12.

Mattock, Karen and Denis Burnham. 2006. Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy* 10(3). 241–265.

Mattock, Karen, Monika Molnar, Linda Polka and Denis Burnham. 2008. The developmental course of lexical tone perception in the first year of life. *Cognition* 106(3). 1367–1381.

Maye, Jessica and LouAnn Gerken. 2000. Learning phonemes without minimal pairs. In S. Catherine Howell, Sarah A. Fish and Thea Keith-Lucas (eds.), *Proceedings of the 24th Boston University Conference on Language Development*, 522–533. Somerville, MA: Cascadilla Press.

Maye, Jessica, Janet F. Werker and LouAnn Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82(3). B101–B111.

Mazuka, Reiko, Yosuke Igarashi and Kenya Nishikawa. 2006. Input for learning Japanese: RIKEN Japanese Mother-Infant conversation corpus. *The Technical Report of the Proceedings of the Institute of Electronics, Information and Communication Engineers* TL 2006-16 (2006-07). 11–15.

Mazuka, Reiko, Yvonne Cao, Emmanuel Dupoux and Anne Christophe. 2011. The development of a phonological illusion: A cross-linguistic study with Japanese and French infants. *Developmental Science* 14(4). 693–699.

Minagawa-Kawai, Yasuyo, Koichi Mori, Nozomi Naoi and Shozo Kojima. 2007. Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *Journal of Neuroscience* 27(2). 315–321.

Miyawaki, Kuniko, Winifred Strange, Robert Verbrugge, Alvin M. Liberman, Jamas J. Jenkins and Osamu Fujimura. 1975. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics* 18(5). 331–340.

Näätänen, Risto, Ann Lehtokoski, Mietta Lennes, Marie Cheour, Minna Huotilainen, Antti Iivonen, Martti Vanio, Paavo Aiku, Risto J. Ilmoniemi, Aavo Luuk, Jüri Allik, Janne Sinkkonen and Kimmo Aiho. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385(6615). 432–434.

Narayan, Chandan 2006. Follow your nose: Non-native nasal-consonant perception in infancy. In David Bamman, Tatiana Magnitskaia and Colleen Zaller (eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development*, 411–422. Somerville, MA: Cascadilla Press.

Nazzi, Thierry, Caroline Floccia and Josiane Bertoncini. 1998. Discrimination of pitch contours by neonates. *Infant Behavior and Development* 21(4). 779–784.

Okazaki, Shuntaro, Shin'ichiro Kanoh, Kana Takaura, Minoru Tsukada and Kotaro Oka. 2006. Change detection and difference detection of tone duration discrimination. *NeuroReport* 17(4). 395–399.

Oller, Kimbrough D. 1973. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* 54(5). 1235–1247.

Peña, Marcela, Atsushi Maki, Damir Kovacic, Ghislaine Dehaene-Lambertz, Hideaki Koizumi, Furio Bouquet and Jacques Mehler. 2003. Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences of United States of America* 100(20). 11702–11705.

Peperkamp, Sharon. 2003. Phonological acquisition: Recent attainments and new challenges. *Language and Speech* 46(2–3). 87–113.

Peperkamp, Sharon and Emmanuel Dupoux. 2002. Coping with phonological variation in early lexical acquisition. In Ingeborg Lasser (ed.), *The process of language acquisition*, 359–385. Berlin: Peter Lang.

Peperkamp, Sharon, Rosenn Le Calvez, Jean-Pierre Nadal and Emmanuel Dupoux. 2006. The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition* 101(3), B31–B41.

Peterson, Gordon E. and Ilse Lehiste. 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32(6). 693–703.

Pickett, Emily R., Sheila E. Blumstein and Mary W. Burton. 1999. Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica* 56(3–4). 135–157.

Pierrehumbert, Janet B. 2003. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46(2–3). 115–154.

Polka, Linda and Janet F. Werker. 1994. Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology. Human Perception and Performance* 20(2). 421–435.

Polka, Linda, Connie Colantonio and Megha Sundara. 2001. A cross-language comparison of /d/-/th/ perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America* 109. 2190–2201.

Port, Robert F., Jonathan Dalby and Micheal O'Dell. 1987. Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America* 81(5). 1574–1585.

Rivera-Gaxiola, Maritza, Juan Silva-Pereyra and Patricia K. Kuhl. 2005. Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science* 8(2). 162–172.

Ross, Elliott D. 1981. The aprosodias. Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology* 38(9). 561–569.

Rose, Yvan and Katherine Demuth. 2006. Vowel epenthesis in loanword adaptation: Representational and phonetic considerations. *Lingua* 116(7). 1112–1139.

Saffran, Jenny R., Janet F. Werker and Lynne A. Werner. 2006. The infant's auditory world: Hearing, speech, and the beginnings of language. In William Damon and Richard M. Lerner (eds.), *Handbook of child psychology*, 58–108. New York: Wiley.

Sakamoto, Tsutomu. 2015. Processing of syntactic and semantic information in the human brain: evidence from ERP studies in Japanese. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Sato, Yutaka, Mahoko Kato and Reiko Mazuka. 2012. Development of single/geminate obstruent discrimination by Japanese infants: Early integration of durational and non-durational cues. *Developmental Psychology* 48(1). 18–34.

Sato, Yutaka, Yuko Sogabe and Reiko Mazuka. 2007. Brain responses in the processing of lexical pitch-accent by Japanese speakers. *NeuroReport* 18. 2001–2004.

Sato, Yutaka, Yuko Sogabe and Reiko Mazuka. 2010a. Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology* 46(1). 106–119.

Sato, Yutaka, Yuko Sogabe and Reiko Mazuka. 2010b. Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *Journal of Cognitive Neuroscience* 22(11). 2503–2513.

Schirmer, Annett and Sonja A. Kotz. 2006. Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences* 10(1). 24–30.

Stager, Christine L. and Janet F. Werker. 1997. Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* 388(6640). 381–382.

Sundara, Megha, Linda Polka and Fred Genesee. 2006. Language-experience facilitates discrimination of /d/-/th/ in monolingual and bilingual acquisition of English. *Cognition* 100(2). 369–388.

Tervaniemi, Mari A., Antti Kujala, Kimmo Alho, Juha Virtanen, Risto J. Ilmoniemi and Risto Näätänen. 1999. Functional specialization of the human auditory cortex in processing phonetic and musical sounds: A magnetoencephalographic (MEG) study. *Neuroimage* 9(3). 330–336.

Tesar, Bruce B. and Paul Smolensky. 1998. Learning optimality-theoretic grammars. *Lingua* 106(1–4). 161–196.

Tesar, Bruce B. and Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.

Tsao, Feng-Ming, Huei-Mei Liu and Patricia K. Kuhl. 2006. Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *Journal of the Acoustical Society of America* 120(4). 2285–2294.

Tsushima, Teruaki, Osamu Takizawa, Midori Sasaki, Satoshi Shiraki, Kanae Nishi, Morio Kohno, Paula Menyuk and Catherine T. Best. 1994. Developmental changes in perceptual discrimination of non-native speech contrasts by Japanese infants: On discrimination of American English /r, l/ and /w, y/. *Technical Report of Institute of Electronic, Information, and Communication Engineers* SP-94-31.

Uffmann, Christian. 2006. Epenthetic vowel quality in loanwords: Empirical and formal issues. *Lingua* 116(7). 1079–1111.

Umeda, Noriko. 1975. Vowel duration in American English. *Journal of the Acoustical Society of America* 58. 434–445.

Vallabha, Gautam K., James L. McClelland, Ferran Pons, Janet F. Werker and Shigeaki Amano. 2007. Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences* 104(33). 13273–13278.

Vance, Timothy J. 1987. *An introduction to Japanese phonology*. Albany, NY: State University of New York Press.

van Lancker, Diana. 1980. Cerebral lateralization of pitch cues in the linguistic signal. *International Journal of Human Communication* 13(2). 227–277.

van Marle, Kristy and Karen Wynn. 2006. Six-month-old infants use analog magnitudes to represent duration. *Developmental Science* 9(5). F41–F49.

Vihman, Marilyn May. 1992. Early syllables and the construction of phonology. In Charles A. Ferguson, Lise Menn and Carol Stoel-Gammon (eds.), *Phonological development: Models, research, implications*, 393–422. Timonium, MD: York Press.

Vouloumanos, Athena, Kent A. Kiehl, Janet F. Werker and Peter F. Liddle. 2001. Detection of sounds in the auditory stream: Event-related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience* 13(7). 994–1005.

Wang, Yue, Joan A. Sereno, Allard Jongmen and Joy Hirsch. 2003. fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience* 15(7). 1019–1027.

Werker, Janet. F, Ferran Pons, Christiane Dietrich, Sachiyo Kajikawa, Laurel Fais and Shigeaki Amano. 2007. Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition* 103(1). 147–162.

Werker, Janet F. and Henny H. Yeung. 2005. Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences* 9(11). 519–527.

Werker, Janet F., John H. Gilbert, Keith Humphrey and Richard C. Tees. 1981. Developmental aspects of cross-language speech perception. *Child Development* 52(1). 349–355.

Werker, Janet F. and Richard C. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7(1). 49–63.

Werker, Janet F. and Richard C. Tees. 1992. The organization and reorganization of human speech perception. *Annual Reviews of Neuroscience* 15. 377–402.

Werker, Janet F. and Richard C. Tees. 1999. Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology* 50. 509–535.

Werker, Janet F. and Richard C. Tees. 2005. Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology* 46(3). 233–251.

Wernicke, Carl. 1874. The symptom complex of aphasia: A psychological study on an anatomical basis. Translated and republished in Robert S. Cohen and Marx W. Wartofsky (eds.), *Boston studies in the philosophy of science* vol. 4., 34–97. Boston: Reidel.

White, Katherine S., Sharon Peperkamp, Cecilia Kirk and James L. Morgan. 2008. Rapid acquisition of phonological alternations by infants. *Cognition* 107. 238–265.

Wong, Patrick C. M., Lawrence M. Parsons, Michael Martinez and Randy L. Diehl. 2004. The role of the insular cortex in pitch pattern perception: The effect of linguistic contexts. *Journal of Neuroscience* 24(41). 9153–9160.

Yeung, Henny H., Heng Ke Chen and Janet F. Werker. 2013. When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language* 68(2). 123–139.

Zatorre, Robert J., Alan C. Evans, Ernst Meyer and Albert Gjedde. 1992. Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256(5058). 846–849.

Zatorre, Robert J. and Jackson T. Gandour. 2008. Neural specializations for speech and pitch: Moving beyond the dichotomies. *Philosophical Transactions of Royal Society B Biological Sciences* 363(1493). 1087–104.

Zatorre, Robert J and Pascal Belin. 2001. Spectral and temporal processing in human auditory cortex. *Cerebral Cortex* 11(10). 946–953.

Zatorre, Robert J., Pascal Belin and Virginia B. Penhune. 2002. Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences* 6. 37–46.

Mutsumi Imai and Junko Kanero

# 2 The nature of the count/mass distinction in Japanese

## 1 Introduction

The count/mass distinction has been noted as one of the most fundamental conceptual distinctions, as it is directly relevant to the identity of entities in the world. Objects are individuated, whereas substances are non-individuated. When we say that two objects are "identical" or "the same", we are referring to "two objects in their entirety", and not to "two distinctive parts of a single object". In contrast, when we talk about the "identity" or "the sameness" of substances, there is no notion of wholeness. Substances are of "scattered existence", and there is no such thing as "whole sand", "whole water", or "whole clay" (cf. Quine 1969). This portion of sand is identical to that portion of sand, as long as the two portions consist of the same physical constituents.

This difference in identity or sameness between objects and substances leads to fundamentally different extension principles for determination of category membership across the two ontological kinds. For example, the label "cup" is applied to any whole object of a similar "cup" shape which can potentially contain liquid, regardless of its color and material components. If a "cup" is broken into pieces, each porcelain piece no longer constitutes a "cup". In contrast, the word "clay" is extended to any portion of clay, irrelevant of shape. One can divide a portion of clay into many small pieces, and each piece is still clay.

### 1.1 The gavagai problem

One extremely important question is how infants come to know this ontological constraint for word learning. Infants must learn meanings of words by inference from a single or very limited number of examples. In so doing, they face a well-known problem of induction, that is, the "gavagai" problem posed by Quine (1960, 1969): When someone goes to a place where he does not know a single word of the language spoken there, he must guess meanings of words, and the only clue he has is what he can observe in the situation in which a given word is uttered. However, it is virtually impossible to determine the referent of the word in the situation, let alone the meaning of it. If the traveler sees a rabbit and hears the word *gavagai* spoken at

the same time, how could he know whether the word refers to the whole white fluffy animal itself, or to an unbounded substance, such as the animal's fur or its meat?

Quine (1960, 1969) argues that this indeterminacy of word meanings is a real problem for learners of any language, using the example of how the English words *ox* and *cattle* should be translated into Japanese. In Japanese, whenever a noun is enumerated, it has to be accompanied by a classifier. For example, English phrase *five oxen* is translated as *go too no usi* (five CLF GEN cattle), in which *too*, a classifier for big animals, functions as a unit of quantification roughly equivalent to the English classifier *head*. Thus, a direct translation of *five oxen* is not possible in Japanese. The closest translation of "five oxen" in Japanese might be "five heads of cattle".

Coming back to the problem faced by infants learning their first language, it is critical that they know whether the word whose meaning they have to infer refers to a bounded thing or to an unbounded thing, because even if an infant could correctly identify the small white animal in front of him as the referent of the word *usagi* 'rabbit', generalization to other referents is only possible when he also knows the ontological principle of word meaning extension – that names for bounded things should be generalized on the basis of the *sameness* of the thing as a whole, where names for unbounded things should be generalized on the basis of the *sameness* of its material constituent.

How do infants come to know this principle before they start learning the meanings of words? One clue they may have is the difference in the word forms which codify bounded things and unbounded things. For example, in English, infants may come to know the form difference between count nouns and mass nouns, and then that the form difference corresponds to a conceptual distinction between bounded things and unbounded things. In fact, Quine (1969) conjectured that this is how (English-speaking) infants acquire the ontological difference between object kinds and substance kinds.

## 1.2 Count/mass distinction in language and its psychological consequences

Quine's conjecture (1969) evoked many questions, and there are at least three directions in which the issue concerning the relation between language and the ontological concept has been addressed. In the first direction, Quine's thesis about the relation between acquisition of count/mass grammar and acquisition of the ontological distinction between objects and substances has been questioned. In the second direction, Quine's assumption that Japanese and/or other classifier languages lack the count/mass distinction has been questioned. The third direction has to do with linguistic relativity. If we take Quine's thesis to an extreme, we must predict that speakers of English-type languages with obligatory count/mass marking and speakers

of Japanese-type languages with classifier grammar must have drastically different world views. The former would think that entities in the world are divided into two kinds of things, those which are bounded and individuated and those which are unbounded and non-individuated, whereas the latter would not care about this distinction or perhaps not even notice it. Researchers have asked whether this hypothesis, which can be formulated as a linguistic relativity – or the Whorfian – hypothesis, is tenable.

## 1.3 Overview of the chapter

This chapter is divided into three sections. In the first part, we address whether Japanese children, whose ambient language does not have transparent and systematic count/mass marking, are able to constrain inference of word meanings by the ontological principles of word meaning generalization. If Quine's conjecture is correct, Japanese infants should not be able to learn the ontology-based extension principles for object names and substance names, because they lack the necessary clues in their linguistic input. We review our previous work examining this question (Imai and Gentner 1997; Imai and Mazuka 2007; cf. Soja, Carey and Spelke 1991). We conclude that Quine is incorrect, and that children DO acquire the ontological distinction even when their language does not have apparent syntactic marking of count nouns and mass nouns.

While finding a cross-linguistically shared appreciation of the ontological distinction, Imai and Gentner (1997) also found a substantial difference across English speakers and Japanese speakers on their construal of boundedness in various types of physical entities (e.g., T-joint pipe). We explore whether the difference between English and Japanese in marking count nouns and mass nouns penetrates into people's notion of "sameness" in a non-linguistic context, and if so, how the cross-linguistic difference arises.

In the second part, we ask whether Quine's assumption (see also Chierchia 1998 and Lucy 1992) – that Japanese and/or other classifier languages do not have a syntactic distinction between count nouns and mass nouns – itself is correct. Several theorists argue against this view and have proposed different analyses of the count/mass status of classifier languages. In particular, some researchers have argued that classifier languages do indeed have count nouns and mass nouns, and the distinction is grammatically realized through the use of distinct classifiers for count nouns and for mass nouns (Chen and Sybesma 1998, 1999; Yi 2009, 2010). We report on an experiment using event-related potentials (ERPs) that empirically tested whether and how the count/mass distinction is processed in Japanese speakers' brains. We did not find the evidence that the count/mass distinction is distinctively made in the brain by the use of classifiers. Instead, the results suggest that processing of a

classifier phrase (a noun + numeral + classifier; e.g., *enpitu ni hon* [pencil two long-thing-CLF] 'two pencils') *in general* is semantic-based although it may involve some syntactic aspects; the relative weight placed on semantic and syntactic processes varied across different types of nouns paired with different types of classifiers, and the neural process seems to rely more on semantic information when object names were paired with mass classifiers (e.g., *enpitu huta sazi* [pencil two spoon] 'two spoons of pencil').

In the final section, we integrate the two parts to draw general conclusions concerning the linguistic representation of object names and substance names in the minds of Japanese speakers and concerning the relation between language and cognition. We close the chapter by addressing unsolved problems and directions for future research.

# 2 Part I: Language and the ontological distinction

## 2.1 Is Japanese children's word learning constrained by the ontological principles?

This section addresses the issue of whether Japanese children are able to constrain the inference of word meanings by the ontological distinction, or whether they get stuck on Quine's (1960, 1969) "gavagai problem" described earlier.

Imai and Gentner (1997) tested this question by comparing Japanese-reared and English-reared children of three age groups (early 2-year-olds, late 2-year-olds, 4-year-olds) and adults. Imai and Gentner devised a word extension task in which the experimenter introduced a novel word (e.g., *dax*) in association with an unfamiliar physical entity that the children had never seen before. Participants were presented with a target entity and taught a label for it. They were also shown two test items and were asked to judge to which of the two alternative entities the label should be applied. One of the test items was the same as the target with respect to shape but different in material. The other alternative entity was the same as the target with respect to material composition but different in shape. A child's choice of either the same-shape or the same-material alternative was considered to reveal which of the two dimensions that child was using for generalizing the novel label.

To minimize the influence from the grammatical construction on the inference of the meaning of the novel label, Imai and Gentner (1997) used specific wordings. For English speakers, the novel words were carefully introduced in such a way that participants could not know whether the entity was syntactically seen as a count or a mass noun, e.g., (1) (*This dax* and *the dax* could be used for either an object or a substance; in contrast to *some dax* which could only be used with a substance).

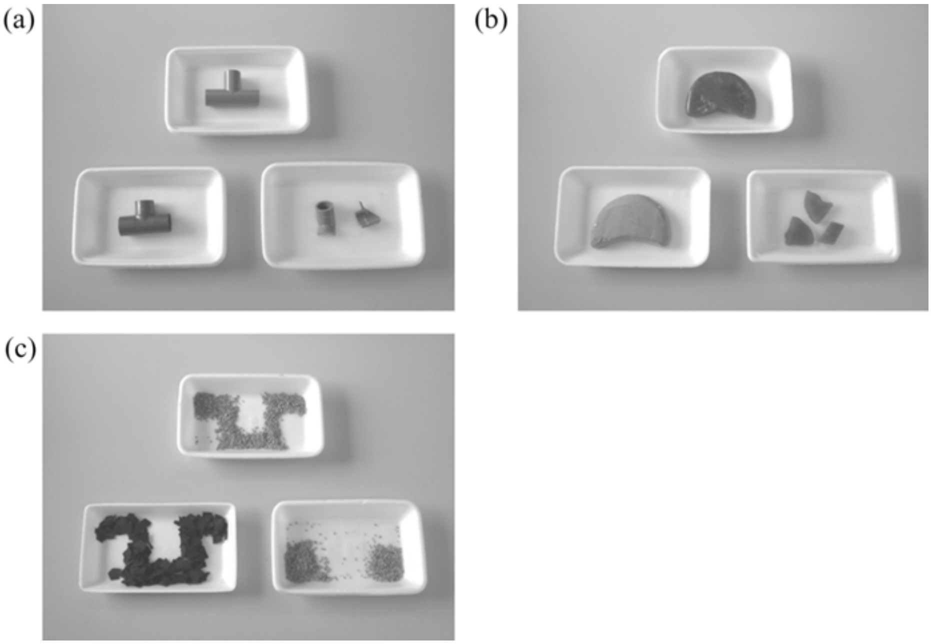



**Figure 1:** Sample material sets for (a) a complex object trial; (b) a simple object trial; (c) a substance trial

(1) *Look at this* dax. *Can you point to the tray that also has the* dax *on it?*

Because the grammatical structure of Japanese does not reveal the noun's status of individuation, sentences in Japanese naturally did not provide countability information about the target entity, e.g., (2). (The use of bare noun such as *dax* is a very natural way to refer to either substances or objects in Japanese.)

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| (2) | *Kore* | *wa* | *dax* | *desu.* | *Dotira* | *no* | *sara* | *ni* | *dax* | *ga* | *aru?* |
| | this | TOP | dax | is | which | GEN | tray | LOC | dax | NOM | exist |

'This is a dax. Which tray is a dax on?'

Imai and Gentner (1997) then set up three different types of physical entities. The first type, the *complex objects*, were real artifact objects that had fairly complex shapes and distinct functions. For example, a T-joint pipe made of plastic (target) was presented along with a metal T-joint pipe (shape test) and broken pieces of the target (material test). If the participant pointed to the metal pipe, it was assumed to be an indicator that he or she construed the target entity as a countable object. In contrast, if the participant pointed to the plastic pieces, it would indicate that he or she saw the target entity as an uncountable substance (Figure 1a). The second type

of entity, the *simple objects*, had very simple structures with no distinct parts (Figure 1b). They were made of a solid substance, such as wax, and were formed into a very simple shape. For example, a kidney-shaped piece of wax (target) was presented together with a kidney-shaped piece of plaster (shape test) and some wax pieces (material test). The third type of entity, the *substances*, were nonsolid substances, such as sand or hair-setting gel, that were arranged into distinct, interesting shapes when presented. For example, a target of wood chips formed into a U-shape was presented together with tiny leather pieces configured into a U-shape (shape test) and piles of wood chips (material test) (Figure 1c). Here, Imai and Gentner hypothesized that solid entities with complex and cohesive structures would be more naturally (and perceptually) individuated than entities with simple structures. They also hypothesized that entities with simple structures would be more naturally individuated than nonsolid substances.

## 2.2 Understanding of the ontological principles in Japanese children

Both Japanese- and English-reared children and adults clearly showed similar classification behavior based on the entities' perceptual appearance. All participants tended to show an object construal and to extend the labels by shape when they engaged in the complex object trials. They were more likely to show a substance construal when they engaged in the substance trials. It seemed that even 2-year-old Japanese children, whose ambient language does not provide a systematic and easily perceptible syntactic distinction across object names and substance names, applied different rules for determining identity for complex objects and for substances, and then extended novel words accordingly.

## 2.3 Cross-linguistic differences in the construal of entities

Although English- and Japanese-reared children both clearly understood the ontological principles for extending object names and substance names, when English and Japanese speakers' classifications were compared within each trial type, there was a marked difference in how English and Japanese speakers construed the simple objects and the substances. For example, in the simple object trials, English speakers treated the simple-shaped discrete entities in the same way as the complex objects and showed a clear object construal bias, whereas Japanese children did not show any systematic tendency in their classification. In fact, Japanese adults tended to see the simple objects more as lumps of uncountable substances, choosing the material alternative more often than the shape alternative. And in the substance trials, whereas Japanese speakers almost always generalized novel words based on

the material identity, English speakers did not show any preference between the shape identity and the material identity.

Imai and Gentner's (1997) results suggested that the ontological distinction between objects (e.g., *pipe*) and substances (e.g., *wax*) is understood even among children whose language has no apparent count/mass syntax. This result refutes the strong version of linguistic relativity (e.g. Quine 1969) and suggests that the ontological distinction is universally shared. At the same time, their results also uncovered noteworthy cross-linguistic differences between the two language groups in a way that was consistent with linguistic relativity (Whorf 1956; see also Lucy 1992).

Lucy (1992), an anthropological linguist, also advanced a linguistic relativity position. Like Quine (1969), he assumed that classifier languages treat all nouns as mass nouns. He predicted that speakers of a classifier language should show a stronger attention to the material constitution of entities than English-type languages which have obligatory count/mass grammar, and that they would tend to construe entities in light of their material composition even outside the realm of language, i.e. when simply perceiving and categorizing things.

Does the cross-linguistic difference found in the word extension task transfer to classification in a non-linguistic context, then? Many studies have reported that children tend to form more adult-like, consistent categories when asked to determine an extension of a novel label (i.e., to find new referents of the label given to a target entity) than when asked to determine the "same" object without using any labels (e.g., Imai et al. 1994; Landau, Smith, and Jones 1988; Markman and Hutchinson 1984; Waxman and Gelman 1986; Waxman and Kosowski 1990). If that is the case, the language effect might be weakened when people engage in a no-word classification task.

To examine this possibility, Imai and Mazuka (2007) tested Japanese-speaking and English-speaking 4-year-olds and adults, using a no-word classification task. The stimuli and the procedure were the same as in the word extension task used by Imai and Gentner (1997), except that word labeling was not involved. The participants were presented with a target entity and two alternatives and were then asked to select which of the alternatives was the same as the target entity. The English instruction was (3a) and the Japanese instruction was (3b).

(3) a. *Show me what's the same as this.*

b. *Kore to onazi no wa dotti desuka.*
this with same one TOP which COP Q
'Which one is the same as this one?'

The results in general indicated that, across the three trial types, Japanese speakers put more weight on the material in determining what would be the same

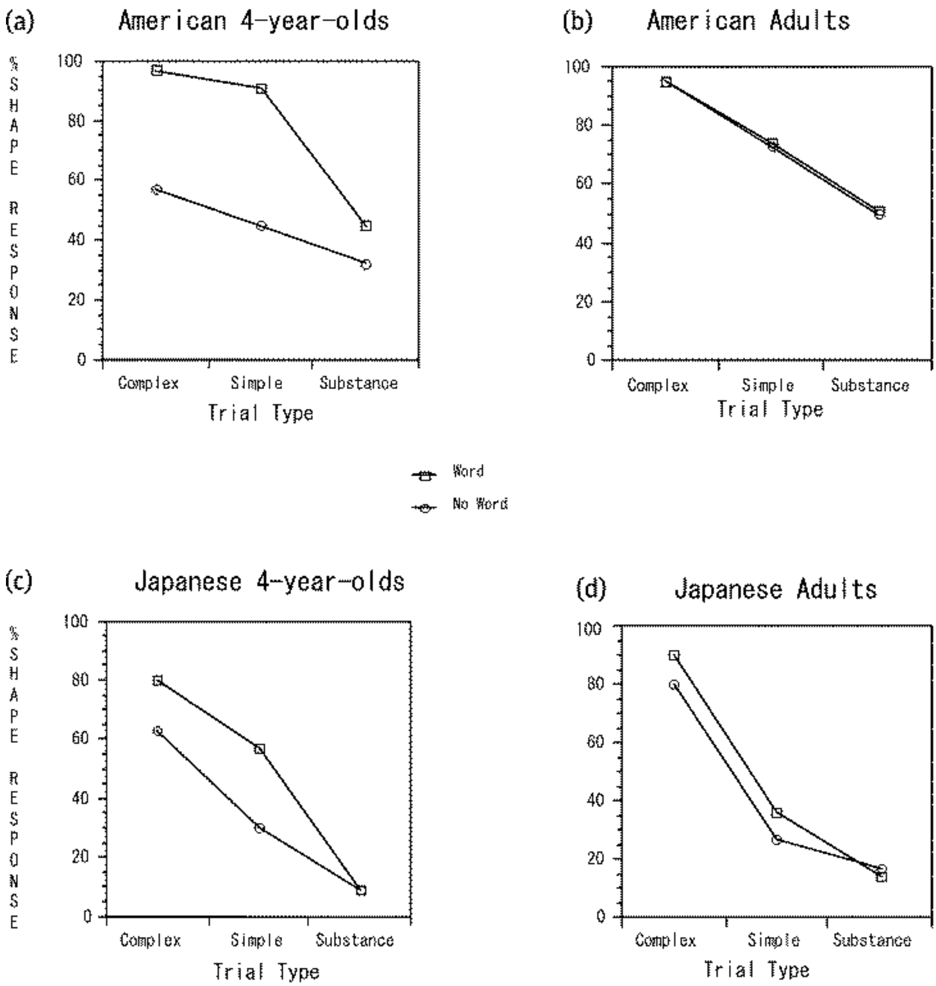



**Figure 2:** Subject's classification behavior in the no-word context in the word extension (neutral-syntax) tasks and non-lexical classification task: (a) American 4-year-olds, (b) American adults, (c) Japanese 4-year-olds, and (d) Japanese adults. (Adopted from Imai and Mazuka 2007)

as the target (material bias), whereas English speakers put more weight on shape (shape bias). Thus, the cross-linguistic difference found in the word extension task (Imai and Gentner 1997) was replicated in the no-word categorization task.

The detailed analysis revealed that the adults' performance in this no-word classification task was virtually identical to that observed in the word extension task, as shown in Figure 2b (English-speaking adults) and Figure 2d (Japanese-speaking adults). In the simple object trials, for example, adult English speakers and adult Japanese speakers showed the opposite classification patterns. But in

contrast to adults, children's classification styles in the no-word classification task were very different from the styles they showed in the word extension task. This discrepancy between the word extension and no-word classification tasks was particularly large in English-speaking children (see Figure 2a). Whereas the English-speaking children in the word extension task showed virtually the same response patterns as the adult English speakers, their performance in the no-word categorization task was at a chance level in all three trial types.

## 2.4 How do language-specific biases arise?

What can be concluded so far from the results of Imai and Gentner's (1997) and Imai and Mazuka's (2007) studies? First, participants' classification between objects and substances is universally constrained by the ontological distinction, regardless of whether the speaker's native language grammatically marks this distinction. However, at the same time, it appears that language-specific syntactic structures can influence the construal of entities (like those used in the simple object trials) that are located around the boundary of the two ontological kinds. The structure of the English language seems to bias English speakers toward the object construal (i.e., there is a bias to classify perceptually ambiguous entities based on shape such as the kidney shape), whereas the structure of the Japanese language seems to bias Japanese speakers toward the substance construal (i.e., there is a bias to classify perceptually ambiguous entities based on material such as wax). Furthermore, it seems that the language-specific construal of entities first becomes evident in the context of word learning, and gradually develops into general default construal that manifests itself without invocation of labels.

How does the shape bias arise in English speakers? Because the count/mass distinction is obligatory, it seems likely that even though Imai and Gentner (1997) presented a novel label in a syntactic frame in which the noun's count/mass status would not be revealed, the English speakers in that study did not encode the noun as having a "neutral" or "indeterminate" syntactic status. Rather, in assigning either count or mass syntactic status to the nouns, the children may have assumed by default that the nouns were count nouns rather than mass nouns, because the count interpretation is more common for "the/this/that X" (cf. Samuelson and Smith 1999).

The results of a control study by Imai and Mazuka (2007, Experiment 3) supported this possibility. In this experiment, the stimuli and the procedure were exactly the same as those used in the Imai and Gentner (1997) study of word extension with ambiguous syntax, with one exception: Each novel noun was presented either in the count noun or the mass noun syntactic frame. The participants in the count noun condition heard novel nouns in the count noun syntax throughout, across the three entity types (complex object trials, simple object trials, and substances trials). Likewise, for those in the mass noun condition, the novel nouns were presented in

the mass noun syntax in all the trials. The instruction used in the count syntax condition was (4a) and the instruction for the mass noun condition was (4b).

(4) a. *Look! This is an X* (pointing the target entity). *Can you point to* another *X?*

b. *Look! This is X. Can you point to* some more *X?*

As shown in Figure 3, when novel nouns were presented in the mass noun syntactic frame, the default classification pattern (i.e., the pattern in the ambiguous syntax case in Imai and Gentner's 1997 study) was drastically changed by the syntactic markers. The English-speaking adults' response pattern in the mass noun condition showed a random response in the complex object trials (48%), presumably because the complex objects invite the object construal very strongly and the syntactic information conflicts with this strong default construal. In contrast, they showed a material bias in the simple object trials (85% material response). This suggests that, despite a strong bias toward construing a simple-shaped solid lump of substance as an individuated object, they were fully capable of mapping a novel label to the material of the entity when prompted to do so by syntactic cues. In the substance trials, again they selected the material alternative highly above chance level (87%).

The response pattern shown by the English-speaking children in the mass noun condition was overall very similar to the adults' pattern, showing a random response pattern in the complex object trials, and a high rate of material responses in the substance trials (59% and 19.6% shape response, respectively). However, in contrast to the adults, the 4-year-olds' performance in the simple object trials was at the chance level (46% shape response). Recall that in the word extension task in which the ambiguous syntax has been used, the English-speaking children's shape response level was very high (91%), in fact almost as high as that for the complex objects (95%). Even though their performance in the simple object trials was still at the chance level with the use of the mass noun syntactic frame, their shape-based responses decreased by 45% from those in the ambiguous syntax case. Therefore, English-speaking 4-year-olds definitely knew that mass noun syntax flags a target entity as a substance (see also Subrahmanyan, Landau and Gelman 1999, for similar findings). However, because they were very strongly biased toward construing any discrete entities as individuated objects (Bloom 1994; Shipley and Shepperson 1990), it must have been difficult for them to overcome this bias and to construe the entities used in the simple object trials as portions of substances.

The English speakers' response pattern in the count syntax condition was intriguing. Their performance here was almost identical to that shown in the ambiguous syntax word extension task, showing a very high rate of shape responses. This is not surprising for the complex and simple object trials, because the rates of shape responses in these two trial types were already at ceiling even in the ambiguous

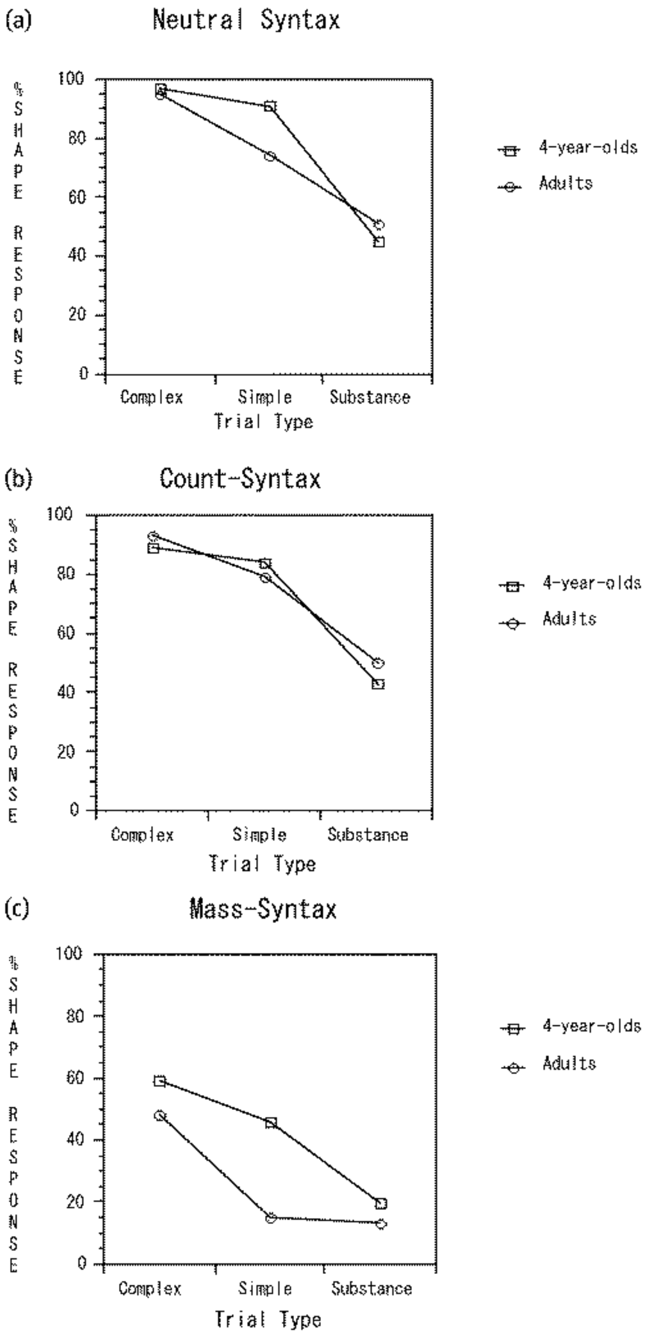



**Figure 3:** English speakers' classification behavior in the (a) neutral-syntax condition (from Imai and Gentner 1997), (b) count-syntax condition, and (c) mass-syntax condition (from Imai and Mazuka 2007)

syntax case. For the substance trials, however, both the children and the adults responded randomly, just as in the ambiguous syntax case.

This pattern supports the idea that English speakers had assumed that the novel nouns presented in the ambiguous syntactic frame were actually count nouns. Together with the result that English-reared children showed the language-specific shape bias just like adults in the word-extension task but not in the no-word classification task, this supports the idea that English speakers' strong shape bias originates in the tendency during word learning in childhood to think that nouns by default are count nouns, and this bias gradually becomes into a general cognitive bias that is applied outside the contexts of label extension.

The degree to which the Japanese speakers' classification was influenced by language is not so clear, because there are two possible ways of interpreting the results. As argued by some linguists and philosophers (e.g., Chierchia 1998; Lucy 1992; Quine 1973), all nouns may indeed be mass nouns in Japanese, and as a consequence, Japanese speakers may develop a strong focus on the material constituent (Lucy 1992). Alternatively, Japanese speakers' understanding of physical entities could be interpreted as a direct reflection of the entity's perceptual nature, in which the classifier markers do not play any important role in object/substance classification.

To further scrutinize this issue, the syntactic status of the count/mass distinction in Japanese needs to be revisited. In the next part, we explore whether any count/mass distinction is marked either *syntactically* or *semantically* in Japanese speakers' minds by examining their event-related potentials (ERPs).

# 3 Part II: Do classifier languages syntactically distinguish count nouns and mass nouns?

## 3.1 Count/mass distinction in classifier languages

While Quine and other theorists maintain that nouns in classifier languages are grammatically mass nouns (the mass noun hypothesis, Chierchia 1998; Lucy 1992; Quine 1960), others propose that numeral classifiers can be categorized into count classifiers and mass classifiers, which are used primarily to specify the amount of objects and substances respectively. Count classifiers are used for bounded objects and provide semantic basis for object classification. In contrast, mass classifiers simply provide units for quantification for unbounded things (usually, but not always, substances).

Cheng and Sybesma (1998, 1999) and Doetjes (1997) claimed that the distinction between the two kinds of classifiers is manifested through differences in their syntactic behavior. That is, classifier languages such as Chinese do mark countability in syntax, not at the level of NP (as is the case with English) but at the level of CLFP (Classifier Phrase).

Mizuguchi (2004) applied a similar view to the Japanese classifier system. According to Mizuguchi, Japanese too has different kinds of classifiers, including both count classifiers and mass classifiers. While count classifiers such as *hon* (a classifier mainly used for long objects) can be used only with countable objects, mass classifiers such as *hai* ('cupful') are used to quantify substances, which must be individuated by containers. See below.

| | | | | |
|---|---|---|---|---|
| (5) | a. | *Pen go hon* | '5 pens' | **Koohii go hon* |
| | b. | *Koohii go hai* | '5 cups of coffee' | **Pen go hai* |

Yi (2009, 2010) also argues that classifier languages have both count nouns and mass nouns, with classifier systems morphosyntactically distinguishing the two. However, unlike the others, Yi proposes *the count noun hypothesis*, as opposed to *the mass noun hypothesis*. Yi's count noun hypothesis asserts that classifier languages have *robust count nouns* which can be syntactically distinguished from mass nouns. For instance, in Chinese, nouns that can take the general classifier *ge* must be count nouns. This rule also applies to Japanese and Korean as they have cognate classifiers, *ko* (or *-tu*) and *kay*, respectively.

Zhang (2012, 2013) proposes yet another view. According to her analysis, the count/mass status of nouns in numeral classifier languages is not binary. Rather, nouns are classified by two features [Numerability] and [Delimitability]. Numerability is the ability of a noun to combine with a numeral directly. Because none of the nouns in classifier languages can take a numeral without an external unitizer, all nouns in such language are [-Numerable]. All nouns in numeral classifier languages are thus non-count nouns. However, non-count nouns are further divided into two classes on the basis of the [Delimitability] feature, which is defined as "the ability of a noun to be modified by a delimitive (size, shape, or boundary) modifier" (Zhang 2012: 1). Nouns that reject a delimitive modifier (e.g., **big sand*, **small water*) are [-Delimitable] and should be considered as mass nouns. Nouns that can be modified by a delimitive modifier (e.g., *big dog*, *big apple*) are *non-mass* nouns.

Zhang (2013) further maintains that classifiers in a numeral classifier language are *unit words*. The [-Delimitable] nouns must be individuated by *individuating* classifiers, which specify a unit of individuation. In Mandarin Chinese, for example, *shui* 'water' is individuated by classifiers such as *di* 'drop', *tan* 'puddle' or *bei* 'cup', which provide an unit for counting, as in *san di/tan/bei shui* (three CLF drop/puddle/cupful of water). In contrast, the [+Delimitable] nouns are individuated by *individual* classifiers, which take the entirety of the referred entity as the unit of individuation (e.g., *san* ***ke*** *xigua* [three CLF-for-fruits watermelon] 'three watermelons'). Importantly, in Zhang's analysis, it is possible to individuate [+Delimitable] nouns by means of individuating classifiers. For example, *watermelon* can be individuated by the individuating classifier *pian* (piece, slice), as in *san* ***pian*** (CLF-pieces) *xigua* (watermelon) 'three pieces of watermelon'. In this case, *xigua* 'watermelon' is treated

as a mass noun. In other words, words like *watermelon* change their count/mass status depending on the type of the classifier that provides the unit for individuation.

In summary, there is no clear agreement on the count/mass status of nouns in classifier languages. Some theorists argue that all nouns in classifier languages are mass nouns, while others argue that nouns in classifier languages also have count/mass status, and that classifiers play a critical role in distinguishing them.

## 3.2 Are Japanese classifiers processed syntactically or semantically? Empirical examinations

Imai and colleagues' work established that Japanese speakers, including 24 month-old infants, do know that different criteria for "sameness" must be applied for object kinds and substance kinds, and that they use this ontological knowledge as a constraint in learning object names and substance names (Imai and Gentner 1997; Imai and Mazuka 2007). Thus, Quine's remark that Japanese speakers do not possess the ontological distinction is incorrect. The critical issue here, then, is whether Japanese speakers represent the two types of nouns (i.e., nouns denoting bounded entities or concepts and nouns denoting unbounded entities or concepts) as syntactically different kinds of words, and if so, whether this grammatical distinction (in any form) influences the acquisition of nouns and the development of the kind of language-specific construal of entities shown by Imai and colleagues.

The choice of classifier does indeed seem to correlate with the count/mass distinction in that different types of classifiers are associated with object names and with substance names. However, whether this distinction is in fact processed syntactically in the speakers' minds is a different issue. As we review in more detail later, it is not clear whether classifiers *in general* are processed syntactically rather than semantically. On one hand, they play a role in syntax, as the lack of a classifier after a numeral clearly produces an ungrammatical sentence, as in (6).

(6) **imooto wa mikan wo* ***ni*** *tabeta*
my young sister TOP orange ACC 2 ate
'My sister ate two oranges.'

At the same time, use of an inappropriate classifier (e.g., *hon*, the classifier for long thin things, as in (7a)) is also sensed anomalous by native speakers of Japanese (Compare the grammatical (7b) in which the numeral *ni* is followed by the classifier for small three-dimensional objects, *ko*). However, it is not clear whether this violation is detected as a semantic violation or a syntactic one.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (7) | a. | **imooto* | *wa* | *2* ***hon*** *no* | *mikan* | *wo* | *tabeta.* |
| | | my young sister | TOP | 2 | orange | ACC | ate |

b. *imooto wa mikan wo ni* ***ko*** *tabeta.*

It is possible that a violation of classifier use across the count/mass distinction (i.e., a count classifier used for a mass noun or a mass classifier used for a count noun) is processed as syntactic violation even if the classifier violation within the count or mass noun category (e.g., use of *hon*, a count classifier, in the place of *mai*, another count classifier) is processed semantically, if count/mass distinction is *syntactically* realized by classifiers (Cheng and Sybesma 1998, 1999; Mizuguchi 2004).

In summary, at present, it is not clear whether count/mass distinction is represented and processed at the level of semantics or at the level of syntax in speakers of a classifier language. One way to empirically investigate this issue is to examine ERPs as classifier phrases are processed in the brain.

## 3.3 Semantic vs. syntactic processing using ERP technique

ERPs are transient electrical signals of the brain that can be observed in response to external stimulation. The identification of an ERP is done by averaging the neural activity elicited by a specific type of stimulation across several dozens of trials. Accumulation of previous research has identified various ERP components, each of which are considered to reliably reflect specific types of neural processing by topographic and temporal characteristics.

The most well-known ERP component in the field of psycholinguistics is the N400. The N400 response is characterized as a negative deflection which is maximal at the centro-parietal region of the scalp that appears approximately 400ms after the onset of the presentation of a semantically anomalous word/phrase. It was first identified by Kutas and Hillyard (1980) with the stimuli of incongruent words within sentences that are semantically anomalous (e.g., *I take coffee with cream and <u>dog</u>*) or improbable (e.g., *He planted string beans in his <u>car</u>*) in the context. Since the first discovery, the N400 has been widely recognized as the neural response to a semantic violation (see Kutas and Federmeier 2011, for a review; see also Sakamoto, this volume).

In contrast, multiple ERP components have been suggested as signatures of syntactic processing. First, the left anterior negativity (LAN) has been identified when morphosyntactic violations such as number disagreement, gender disagreement, and verb inflection violations are detected (e.g., Osterhout and Mobley 1995). The temporal pattern of the LAN often overlaps that of the N400. However, the two components are topographically distinctive. The N400 is most pronounced at the center to posterior region of the scalp whereas the LAN is prominent at the left anterior region.

Another signature associated with syntactic violation, the early left anterior negativity (ELAN), is topographically similar to the LAN, but the ELAN appears earlier, around 150–350ms (e.g., Friederici, Pfeifer and Hahne 1993; Neville, Nicol, Barss,

Forster and Garrett 1991; Hahne and Friederici 1999). The ELAN is most frequently associated with violations of word category or phrase structure rules (e.g., *Der Freund wurde im besucht*. 'The friend was in the visited.'; Friederici et al. 1993).

Lastly, another widely reported ERP component is the P600, which is a relatively long-lasting positivity that appears about 600ms after the onset of the target stimulus (see Hagoort et al. 1999, for a review). This component was first reported as another index of syntactic processing; however, now it is most commonly recognized as the indication of a more general process of reanalysis because the P600 can be elicited by both semantic and syntactic violations. For example, Osterhout and Holcomb (1992) first reported the P600 with incorrect use of transitive verbs, e.g., (8a). However, Van Herten, Kolk and Chwilla (2005) later observed similar effects with semantic anomalies in Dutch sentences, e.g., (8b).

(8) a. *The woman persuaded to answer the door.*

b. *De vos die op de stropers joeg sloop door het bos.*
the fox[sg] that at the poacher[sg] hunted[sg] stalked through the wood
'The fox that hunted the poachers stalked through the woods.'

Thus, although P600 may not be a reliable indicator on its own, it could be used together with other ERP components – N400, LAN, ELAN – to examine the underlying cognitive activities for the words, phrases or sentences in question.

## 3.4 Previous ERP research on noun categorization systems

ERP research has investigated how the brain handles other noun categorization systems such as gender grammar. According to Barber and Carreiras (2005), grammatical gender disagreement between articles and nouns in Spanish sentences evoked the LAN and P600. This LAN-then-P600 pattern was also found in other studies (Gunter, Friederici and Schriefers 2000; Barber, Salillas and Carreiras 2004). Similarly, using German word pairs made up from articles/pronouns and nouns/verbs that are matched (9a) and mismatched (9b), Münte and Heinze (1994) observed a frontally distributed negativity associated with the violation of the gender grammar. The authors concluded that the observed negativity reflects syntactic processing (i.e., LAN), suggesting that grammatical gender is processed primarily on a syntactic basis.

(9) a. *Das-Haus*
The (neutral)-House

b. **Der-Haus*
The (masculine)-House

Although a few studies failed to observe LAN effects (Barber and Carreiras 2003; Hagoort and Brown 1999), the current general consensus of the field seems to be that gender agreement elicits syntactic ERP components except for a few special cases (see Barber and Carreiras 2003).

A few ERP studies have examined the neural processing of Japanese numeral classifiers, although they did not address the count/mass distinction in the use of classifiers. Mueller and his colleagues (2005) investigated neural responses to classifier violation in auditorily presented sentences, e.g., (10a). The violation in this case is the use of the bird-classifier *wa* to count cats. Their experiment tested two other types of violation, i.e., word category violation and case violation. In the word category violation condition, the sentence has a missing noun, resulting in an impossible syntactic phrase structure in which the particle *no* is followed directly by a verb rather than by the expected noun (10b), where a noun is missing between the particle *no* and the verb *tobikoeru*. The case violation involved the misuse of a case-marking particle, such as the use of a nominative particle in which an accusative one would be required, e.g., (10c), where the accusative particle *wo* should follow the second noun phrase *ni hiki no neko*.

(10) a. *Iti wa no kamo ga ni wa no *neko*
one bird-CLF GEN duck NOM two bird-CLF GEN cat
*wo tobikoeru tokoro desu.*
ACC jump over about to COP
'A duck is about to jump over two cats.'

b. *Iti wa no kamo ga ni hiki no*
one bird-CLF GEN duck NOM two small-animal-CLF GEN
**tobikoeru tokoro desu.*
jump over about to COP

c. *Iti wa no kamo ga ni hiki no*
one bird-CLF GEN duck NOM two small-animal-CLF GEN
*neko *ga tobikoeru tokoro desu.*
cat NOM jump over about to COP

Mueller et al. (2005) found that the classifier violation elicited the negativity with a left frontal lateralization. The word category violation and the case violation also resulted in the negativity but it was not left lateralized. The authors thus interpreted that the negativity observed in the classifier violation condition was the LAN, and claimed that Japanese classifiers are processed syntactically rather than semantically.

Sakai and her colleagues (2006), however, found different results. Their study examined the ERPs elicited by visually presented word pairs of a noun and a classifier. In contrast to congruent pairs (11a)–(11c), incongruent pairs of a noun and a

classifier (11d)–(11f) showed a strong negativity around 250–550ms after the onset of the presentation of the classifier.

(11) a. *enpitu* *san* *bon*
pencil three long-object-CLF

b. *tomodati* *san* *nin*
friend three human-CLF

c. *kami* *san* *mai*
paper three flat-object-CLF

d. *ki* *san* *nin*
tree three human-CLF

e. *sensei* *san* *ko*
teacher three small-object-CLF

f. *megusuri* *san* *tyaku*
eye drop three clothes-CLF

As the negativity was not lateralized to the left side of the scalp, the authors interpreted the response to be the N400 and concluded that Japanese numeral classifiers are processed at the semantic level.

Whereas Mueller et al. (2005) argued that the neural processing of classifiers is primarily syntactic-based, Sakai et al. (2006) maintained that it is semantic-based. Thus, it is difficult to draw a clear conclusion as to whether the brain treats classifier violation as syntactic violation or semantic violation. One easy way to resolve this discrepancy is to assume that either or both studies misinterpreted their ERP data. The negativity found by Sakai et al. (2006) may not be the N400 as it was shifted more to the front part of the scalp than the typical N400. On the other hand, the negative deflection observed by Mueller et al. (2005) may not be the LAN, as it lasted much longer than typical LAN effects. The prolonged negativity found by Mueller et al. may reflect increased working memory load rather than syntactic processing (e.g., Marítn-Loeches et al. 2005; Yasunaga and Sakamoto 2007).

Thus, the currently available data on violations in classifier phrases are somewhat difficult to interpret. Yet there is a third possibility. The two different results by Mueller et al. (2005) and Sakai et al. (2006) may be telling us that the Japanese classifier system is a semantically-oriented grammatical system. Classifiers are grammatical morphemes that must accompany nouns with numerals; however at the same time, classifiers semantically classify nouns. This differs from the case of grammatical gender, in which assignment of gender class to each noun is in most cases semantically arbitrary (i.e., gender assignment does not reflect the biological sex of the referent of the noun).

In summary, research on whether processing of classifiers in general recruits a semantic network or a syntactic network, or both, has been inconclusive, as is also the case for whether classifiers make a count/mass distinction grammatically. The neural processing of classifiers may involve both syntactic and semantic processes, and a classifier violation for a given noun may elicit both semantic and syntactic ERP signatures regardless of the noun's count/mass status. However, the relative weight on the semantic and syntactic components may differ across different types of classifier violation. In particular, if the count/mass distinction is realized by count classifiers (or individual classifiers) and mass classifiers (or individuating classifiers), then noun-classifier mismatches that go across the ontological boundary (i.e., an object name combined with a mass classifier, or a substance name combined with a count classifier) may invoke stronger syntactic responses in ERPs as compared to those made within the count/mass category boundary (an object name combined with an inappropriate count classifier, or a substance name combined with an inappropriate mass classifier).

## 3.5 Our experiment

To investigate whether speakers of a classifier language syntactically process the count/mass distinction at CLFP (classifier phrase), we conducted an ERP experiment (Kanero, Imai, Hoshio and Okada submitted). This experiment was similar to Sakai et al. (2006) in that it examined the ERP responses to noun-classifier violations using a word-pair paradigm, but we contrasted noun-classifier mismatch violations within and across the count-mass boundary.

### 3.5.1 Stimuli and procedure

The experiment consisted of four different conditions, including a matched (control) condition and three violation conditions: a violation within the count/mass category, a violation across the count/mass category boundary, and an animal-non-animal violation condition. Half of the nouns were names of objects and the other half were names of substances. The same nouns were used across the four conditions, but were accompanied with different classifiers according to the conditions. Thus, each object name or substance name appeared four times through the whole experiment. Classifiers cannot be used without a numeral, so all classifiers were embedded into phrases by adding the number "two" (e.g., *ni hon* [two long-thing-CL]).

We included the animal-non-animal violation condition, in which object/substance names were paired with classifiers for animals, as a separate violation condition. Animals are never counted by classifiers that are associated with non-animates, and non-animate objects are never counted by animal classifiers. We thus compared

the magnitude of the brain response for the within or between count/mass category violations to that for this very clear and strong violation case.

In the within-count/mass-category violation, an object name was followed by an incongruent count (individual) classifier or a substance name was followed by an incongruent mass (individuating) classifier. In the across-count/mass-category violation, on the other hand, an object name was followed by an incongruent mass classifier whereas a substance name was followed by an incongruent count classifier. For instance, the object name *hude* 'brush' was paired with *ken* (object classifiers for houses and other buildings) in the within-category-violation condition (12b), *kire* 'a piece/slice of' in the across-category-violation condition (12c), and *hiki* (small animal classifier) in the animal-non-animal violation condition (12d). Likewise, the substance name *sio* 'salt' was paired with *kire* in the within-category-violation condition (12f), *dai* (classifier for machines and functional artifacts) in the across-category-violation condition (12g), and *wa* (bird classifier) in the animal-non-animal violation condition (12h). In the matched condition, the target object and substance nouns were accompanied by their proper classifiers (12a) and (12e).

(12) a. object matched condition: *hude ni hon*
b. object within-count-category violation: *hude ni ken*
c. object across-count-category violation: *hude ni kire*
d. object animal-non-animal violation: *hude ni hik*i
e. substance matched condition: *sio huta sazi*
f. substance within-mass-category violation: *sio huta kire*
g. substance across-mass-category violation: *sio ni dai*
h. substance animal-non-animal violation: *sio ni wa*

The pool of nouns and classifiers was drawn from commonly used vocabulary, and different combinations of nouns and classifiers were created. Thirty native Japanese adults who did not participate in the main ERP experiment rated the degree of match for each noun-classifier pair on 1–5 scales (1: Highly mismatched, 2: Somewhat mismatched, 3: Neither matched nor mismatched, 4: Somewhat matched, 5: Highly matched). The main ERP experiment used noun-classifier pairs that were rated as highly matched (mean rating score: 4.60) and pairs that were rated as highly mismatched (mean rating score: 1.31 for the within-count/mass-category violation, 1.13 for the across-count/mass-category violation, and 1.04 for the animal-non-animal violation).

A noun was considered as an "object noun" when the entity denoted by the noun has a clear boundary, and its identity is lost when it is broken into pieces, e.g., (13a). "Substance noun" was defined such that the denoted entity had no boundary and passes the universal grinder test suggested by Pelletier (1979), e.g., (13b). Classification of classifiers as "count classifiers" and "mass classifiers" was not so straightforward, as different researchers have used somewhat different criteria

for this classification (Zhang 2013). We considered classifiers whose unit of counting coincides with the whole of an object, i.e., individual classifiers by Zhang's term (14a) (mentioned above) to be "count classifiers". Classifiers that provide a unit of counting for a segment of a thing (14b) or liquid (14c) are used as "mass classifiers".

(13) a. object noun: *koppu* 'cup', *zitensya* 'bicycle'
b. substance noun: *abura* 'oil', *hatimitu* 'honey'

(14) a. count classifiers: *hon, mai, dai*
b. mass classifiers (thing): *kire* 'piece/slice', *kakera* 'piece/chunk', *katamari* 'lump'
c. mass classifiers (liquid): *hai* 'cup', *hukuro* 'package', *bin* 'bottle'

Nouns and classifier phrases were consecutively presented on the monitor each for 800ms. Twenty five participants, all native speakers of Japanese, were asked to press a "correct" or "incorrect" button to indicate if the classifier was appropriate to count the noun. EEGs were recorded from 32 electrodes, and we examined the change in EEG signals after the presentation of classifier phrases in contrast to the 100ms pre-stimulus recording period.

### 3.5.2 Results

Similar to Sakai et al. (2006), we in general found a negative deflection around 300–500ms after the onset of classifier presentation in all three violation conditions both for the object nouns and the substance nouns. The effect was widespread but the strongest at the center of the scalp, as shown in Figures 4a and 4b. A slight shift of the pattern toward the anterior region was observed for all violation conditions except for the across-category violation between object noun and mass classifier (see below for the discussion of this result). Thus, the observed negativity was not a typical N400, for which the greatest deflection is found in the central-parietal part of the scalp, nor was it a typical LAN as the effect was not limited to the left anterior region. However, as the lateralization was very weak, the negativity seems to be more similar to a N400. Alternatively, the negative deflection may reflect both N400 and LAN. Some researchers suggest that, when semantic and morphosyntactic processes simultaneously take place, a hybrid of N400 and LAN effects can be observed (de Vega, Urrutia and Dominguez 2010; Thierry, Cardebat and Demonet 2003). Although it may be difficult to conclude whether the negativity observed in the study is a N400 or a N400-LAN hybrid, our results clearly suggest that the violation of noun-classifier matching largely elicits the same ERP responses regardless of whether the violation went across the count/mass boundary.

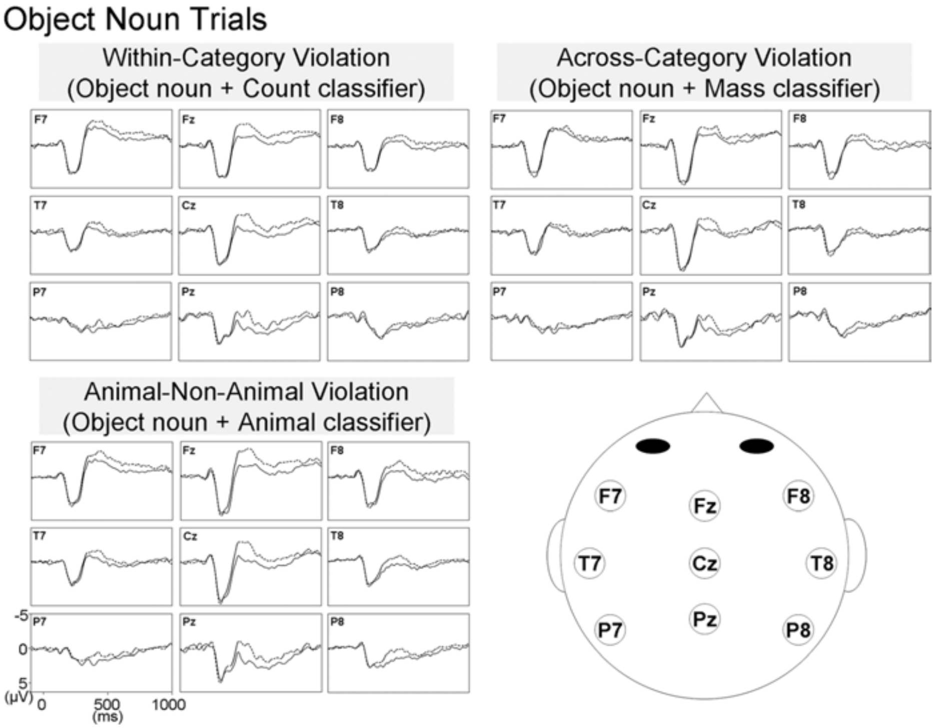



**Figure 4a:** Grand average ERP waveforms of the object noun trails for the within-count/mass-category, across-count/mass-category and animal-non-animal violation conditions (dotted lines) in contrast to the matched condition (solid lines)

Importantly, the incongruent pairs of non-animal object nouns and animal classifiers resulted in larger effects than the within- or across-count/mass-category violation pairs (see Figure 4a). This result demonstrated that the clear separation of animal and non-animal classifiers in the classifier system in Japanese invokes the strongest effects when processing classifier phrases.

We found a topologically unique effect when object nouns were paired with mismatching mass classifiers. As stated earlier, the noun-classifier mismatch pairs except the object noun-mass classifier mismatch pairs elicited an N400-LAN hybrid-like response, which suggests that the matching process of noun and classifier may require both semantic and syntactic bases. However, the object noun-mass classifier mismatch pairs invoked a somewhat different response. In this case, the negative deflection from the baseline was not shifted to the anterior region, which appears to reflect less involvement of syntactic processing. This unique topographical effect may have arisen from the fact that mass classifiers *can* be used to count a wider variety of nouns, even including objects, than count classifiers can. For instance, the classifier phrases such as (15a) and (15b) are unusual, but it can possibly be

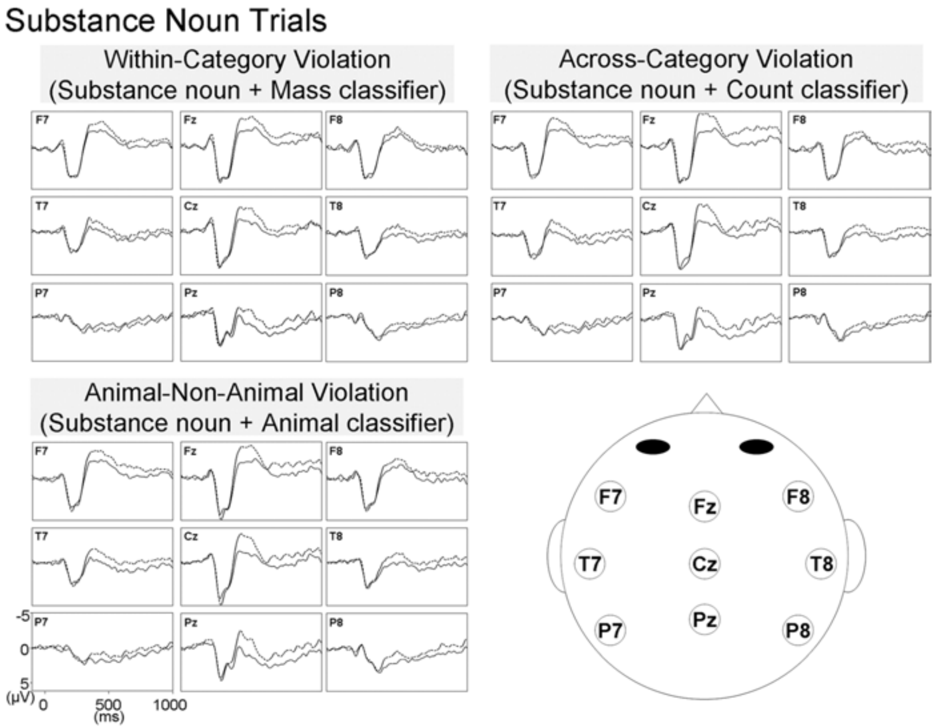



**Figure 4b:** Grand average ERP waveforms of the substance noun trails for the within-count/mass-category, across-count/mass-category and animal-non-animal violation conditions (dotted lines) in contrast to the matched condition (solid lines)

considered as an acceptable phrase if the participants visualize "pencils in two cups".

(15) a. *enpitu* *ni* *hai*
pencil two CLF-cup

b. *kuruma* *hito* *katamari*
car one CLF-pile

In general, when a noun was followed by a mass classifier, a definite mismatch judgment may be difficult to make, requiring greater explorations. For example, even though the phrase "one chunk of pencils" make strike us as anomalous, we may still attempt to form a mental image in which pencils are stuck together to form a chunk. This deeper semantic exploration may have invoked the traditional centroparieto-distributed N400-like effect. In contrast, when a noun-classifier mismatch relies more on memory templates, as is the case for the use of animal classifiers for non-animal objects or substances, syntactic processing may need to be involved.

### 3.5.3 Discussion of the study

The goal of our ERP experiment was to empirically test the proposal that the count/mass distinction is syntactically made by classifiers in numeral classifier languages. We hypothesized that if the count/mass distinction is realized by the classifier phrase, and if this linguistic device bears psychological reality, this distinction would be revealed in ERP responses. Specifically, we expected to see a greater LAN effect when the noun-classifier mismatch went across the count/mass boundary than when the mismatch occurred within the count/mass boundary.

The results did not support this hypothesis. The negative deflection was similarly observed in all mismatched conditions, except for the object noun-mass classifier condition. This suggests that the process of matching a noun and a classifier is done primarily on a semantic basis (with possible additional involvement of syntactic processing) regardless of whether the mismatch goes across the count/mass boundary or not. The count/mass distinction did not affect the magnitude of the response. Instead, the regularity in noun-classifier pairing influenced the magnitude of the response: Animal classifiers are never associated with non-animal object nouns, and thus when this constraint was violated, it invoked the strongest effect even though the noun-classifier mismatches were made within the count/mass boundary.

The fact that the pairing of an object noun and a mass classifier elicited a different ERP response from other types of noun-classifier mismatch is also intriguing. It is important to note that the rated acceptability of the object noun-mass classifier pairs was as poor as the other types of mismatch pairs on the behavioral rating test. However, unlike count classifiers, mass classifiers can in principle be used for individuating objects, as discussed earlier. Thus, participants in the ERP experiment may have searched for a context in which the mass classifier can be used with the object, and this may have caused deeper semantic processing.

Considering this, true noun-classifier mismatch violations going across the count/mass boundary may be limited to cases in which a substance noun is paired with a count classifier, as in *mizu iti dai* (water one CLF-for-machines-and-functional-artifacts). If this is the case, the contrast between the substance noun-count classifier mismatch condition and the substance noun-mass classifier mismatch condition should be critical, and we should expect a larger LAN effect in the former than in the latter condition. However, we did not find a difference in the two cases.

Taken together, the results of the experiments are summarized as follows: (i) Classifier phrases seem to primarily recruit semantic processing, with possible involvement of syntactic processing, and invoke both N400-LAN hybrid-like responses and N400-like responses, depending on the type of noun-classifier mismatch; (ii) When the noun-classifier mismatch can be determined without deep semantic exploration, syntactic processes become prominent, causing a slight shift in topography; (iii) When the noun-classifier mismatch detection is difficult without deep semantic

exploration, semantic processing may become more prominent and a traditional N400-like signature arises.

We thus did not find evidence for the proposal that the count/mass distinction is realized at a syntactic Classifier Phrase level in classifier languages (Cheng and Sybesma 1998, 1999; Mizuguchi 2004), at least in Japanese speakers. However, this conclusion does not necessarily lead to the more general conclusion that the count/mass distinction is purely semantically based. As reviewed earlier, Zhang (2013) proposes that nouns that reject a delimitive modifier (e.g., **big sand*, **small water*) are mass nouns and all other nouns are non-mass nouns in Chinese. It would be interesting to examine whether adjective phrases like **big sand* or **small water* elicit the LAN or N400 in classifier languages. If this is in fact the case, it would indicate that the count/mass distinction (or mass/non-mass distinction) is made syntactically in NP but not CLFP. This needs to be tested in future research.

# 4 Conclusions and future research

This chapter has explored if and how adult Japanese speakers and Japanese-reared infants and young children represent the count/mass distinction. The ontological distinction between object kinds and substance kinds is appreciated and is used as a constraint for word learning by Japanese-reared children at 24 months of age. Thus, it is likely that children come to appreciate the ontological distinction very early even when their native language does not grammatically mark the distinction, and in this sense, syntactic bootstrapping is not necessary for children to acquire object names and substance names.

One may be concerned that it is too early to conclude that classifiers do not contribute to children's awareness of the ontological difference between objects and substances. In fact, our results suggest that the neural processing of mass classifiers is distinctive. It is thus theoretically possible to think that young children use classifiers to acquire an awareness of the count/mass distinction: children may know that, whereas object names usually each have only one matching classifier, substance names can be paired with various classifiers depending on the situation. Making use of the probability of co-occurrence of particular nouns and particular classifiers, they may realize that a noun which is always accompanied by a single classifier belongs to the category of object kinds while a noun which is accompanied by various classifiers across different situations belongs to the category of substance kinds.

This possibility would be supported if children's awareness of the distinction between objects and substances changes after they master the use of classifiers. However, we do not think that this possibility is likely. The acquisition of classifiers is much slower than the acquisition of nouns, and Japanese children as old as age 5½ are still not fully familiar with the use of classifiers (Uchida and Imai 1999).

Considering this, it is more probable that the direction is the other way around: The ontological distinction, which exists from an early age, is used to learn the matching between nouns and classifiers (see Sato and Haryu 2006 for some evidence for this possibility).

Of course, it is possible that children do use classifier information *after* acquiring the meanings of classifiers for inference of noun meanings. For example, when hearing a novel noun *neke* with the classifier *teki* ('drop') in a context where no other cues are available, the child may infer that *neke* is a kind of liquid if she knows the meaning of this classifier. Chien, Lust and Chiang (2003) in fact argue that classifiers could bootstrap noun learning. However, in our view, classifier information, if used at all, can serve only as a weak and secondary cue, which is used in rare situations when the classifier at hand is specific and known by the child but perceptual and social cues are unavailable (Brandone et al. 2007). In actuality, the classifiers children frequently hear in everyday settings such as (16a) are very broad in their applications, and do not sufficiently constrain the meaning of the noun associated with them. More specific, category-based classifiers such as (16b) are usually acquired only through formal school education, so by the time children learn them, they are likely to have already learned the nouns for which these classifiers are used.

(16) a. *ko* (for small three-dimensional objects), *hon* (for long and thin things), *mai* (for flat things), *hiki* (for small animals), *tou* (for large or important animals), *dai* (for machines and some functional artifacts)

b. *satu* (for books and other bounded reading materials), *wa* (for birds), *sao* (for clothes cabinets)

Even though classifiers are not useful for the purpose of constraining the meaning of nouns, asking whether the ontological distinction between object nouns and substance nouns is syntactically processed by adult Japanese speakers is still worthwhile. The results of our experiment suggest that count classifiers and mass classifiers are processed differently. When the participants were simply matching a noun and a classifier, the classifier phrase seemed to be processed both semantically and syntactically. However, to match object nouns and mass classifiers, a more flexible approach is required and thus semantic nature of classifiers seems to be considered more heavily.

These results provide insights onto the issue of how we should characterize grammatical categorization systems in general. Traditionally, researchers tend to want to draw a clear distinction between semantics and syntax. However, the way the classifier system is processed in the brain suggests that such a binary categorization does not properly reflect reality, and that we should explore how semantics and syntax are integrated in the brain.

Of course, Japanese is not the only language with a numeral classifier system, and numeral classifier languages are not homogeneous in their syntactic roles in natural discourse. Brain responses may well reflect any differences which may exist among different numeral classifier languages. For example, Chinese classifiers must be used not only in numeral phrases (e.g., [numeral + classifier] table) but also in phrases with demonstratives (e.g., this [numeral + classifier] table) (the numeral after the demonstrative is often dropped, however). In contrast, Japanese classifiers are only used with numerals and are not used in constructions with demonstratives. Furthermore, in Chinese, a classifier functions as a rough equivalent to an indefinite article, while in Japanese, classifiers are only used when it is pragmatically important to specify the number of things in discourse. For example, as an equivalent to the English phrase: *I have a cat*, Chinese speakers are most likely to say (17a). In contrast, Japanese speakers are most likely to say (17b).

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (17) | a. | *wo* | *yang* | *yi* | *zhi* | *mao.* | |
| | | I | raise | one | small-animal-CLF | cat | |
| | b. | *watasi* | *wa* | *neko* | *wo* | *katte-iru.* | |
| | | I | TOP | cat | ACC | raise-state | |
| | | 'I have a cat/some cats.' | | | | | |

Here, the information "one" is not verbalized by Japanese speakers unless this information is pragmatically important, e.g., when saying: *I have only one cat, but not two*, in response to the question: *Do you have two cats?* It would be expected that these differences would result in a much higher frequency of classifier use in Chinese than in Japanese.

To confirm this intuition, Saalbach and Imai (2012) compared the frequency of classifier use in a Japanese novel and in its Chinese translation, using the Chinese-Japanese parallel corpus[1] of the novel *Bottyan* (Master Daring) written by Soseki Natsume (1906/1964). In the original Japanese text, there were 111 instances of classifiers, while in the Chinese translation, there were 405 instances. Thus, 294 classifier tokens were added in the course of translating the original Japanese text to Chinese. On closer examination, there were 58 cases in which a classifier was used with "one" (*iti*) in the Japanese original. In the Chinese translation, there were 156 cases of "one" (*yi*) with a classifier construction. When the number was "two" or "three", there were 21 classifier instances in Japanese and 53 in Chinese. In the Chinese translation, 175 classifier instances were of the "demonstrative + classifier + noun" construction (e.g., *Zhe zhang weirenzhuang* [this CLF document]). However, in the original Japanese text, all these cases were simple "demonstrative + noun" constructions without a classifier. This study thus revealed that classifiers are used

**1** *Chinese-Japanese parallel corpus* by the Beijing Center for Japanese Studies was used.

roughly four times as often in Chinese as in Japanese, which is consistent with our structural analysis of the Chinese and Japanese classifier systems.

Given this linguistic difference between Japanese and Chinese, it is extremely important to extend our ERP study to Chinese and other numeral classifier languages. Close examination of the way the semantic and syntactic nature of nouns and classifiers in a given classifier language correlates with ERP responses in the speakers' brain in different numeral classifier languages will allow us to understand whether and how the count/mass distinction is represented in the minds of speakers of a classifier language, which in turn might further help us understand the interaction between semantic and syntactic processing.

The lack of grammatical expression of the count/mass distinction and the use of a numeral classifier system are prominent features that make Japanese distinctive from many other languages such as English. The uniqueness of the Japanese classifier system enriches research on how Japanese treats a range of fundamental concepts such as number, animacy, and the object/substance distinction. Further, it is a window into the bigger picture of how language systems and the ontological understanding of the world are interrelated. Developmental and neurophysiological research suggests that the numeral classifier system does not serve as a primary basis for the ontological object/substance distinction in Japanese speakers. However, it also demonstrates that Japanese- and English-speakers rely on different cues to judge "sameness" of entities, suggesting some influence of having the classifier system to non-linguistic concepts. Our ERP study indicates that the Japanese classifier system is not a set of strictly grammatical systems but a complex system that integrates both semantic and syntactic information. Future research including examination of other classifier languages is needed to further reveal the universal nature of classifier systems as well as the unique nature of the Japanese classifier system.

## References

Barber, Horacio and Manuel Carreiras. 2003. Integrating gender and number information in Spanish word pairs: An ERP study. *Cortex* 39(3). 465–482.

Barber, Horacio and Manuel Carreiras. 2005. Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience* 17(1). 137–153.

Barber, Horacio, Elena Salillas and Manuel Carreiras. 2004. Gender or genders agreement? In Manuel Carreiras and Charles Clifton (eds.), *On-line study of sentence comprehension: Eye-tracking, ERP and beyond*. 309–327. Brighton, UK: Psychology Press.

Bloom, Paul. 1994. Possible names: The role of syntax-semantics mappings in the acquisition of nominals. *Lingua* 92. 297–329.

Brandone, Amanda C., Khara Pence, Roberta M. Golinkoff and Kathy Hirsh-Pasek. 2007. Action speaks louder than words: Young children differentially weight perceptual, social, and linguistic cues to learn verbs. *Child Development* 78(4). 1322–1342.

Cheng, Lisa Lai-Shen and Sybesma Rint. 1998. *Yi-wan tang, yi-ge tang: Dan wei ci yu wu hua ci*. [Yi-wan tang, yi-ge tang: Classifiers and massifiers.] *Tsing-Hua Journal of Chinese Studies* 28(3). 385–412.

Cheng, Lisa Lai-Shen and Sybesma Rint. 1999. Bare and not so bare nouns and the structure of NP. *Linguistic Inquiry* 30(4). 509–542.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6(4). 339–405.

Chien, Yu-Chin, Barbara Lust and Chi-Pang Chiang. 2003. Chinese children's comprehension of count-classifiers and mass-classifiers. *Journal of East Asian Linguistics* 12(2). 91–120.

de Vega, Manuel, Mabel Urrutia and Alberto Dominguez. 2010. Tracking lexical and syntactic processes of verb morphology with ERP. *Journal of Neurolinguistics* 23(4), 400–415.

Doetjes, Jenny. 1997. *Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch and English*. Hague, HAG: Leiden University dissertation.

Friederici, Angela D., Erdmut Pfeifer and Anja Hahne. 1993. Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research* 1(3). 183–192.

Gunter, Thomas C., Angela D. Friederici and Herbert Schriefers. 2000. Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience* 12(4). 556–568.

Hagoort, Peter and Colin M. Brown. 1999. Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research* 28(6). 715–728.

Hagoort, Peter, Colin M. Brown and Lee Osterhout. 1999. The neurocognition of syntactic processing. In Colin M. Brown and Peter Hagoort (eds.), *The neurocognition of language*, 273–316. Oxford, UK: Oxford University Press.

Hahne, Anja and Angela D. Friederici. 1999. Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience* 11(2). 194–205.

Imai, Mutsumi and Dedre Gentner. 1997. A crosslinguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition* 62(2). 169–200.

Imai, Mutsumi, Dedre Gentner and Nobuko Uchida. 1994. Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development* 9(1). 45–75.

Imai, Mutsumi and Reiko Mazuka. 2007. Revisiting language universals and linguistic relativity: Language-relative construal of individuation constrained by universal ontology. *Cognitive Science* 31(3). 385–414.

Kanero, Junko, Mutsumi Imai, Noriko Hoshino and Hiroyuki Okada. In submission. Is count/mass distinction in Japanese syntactic or semantic?: Insights from Event Related Potentials.

Kutas, Marta and Kara D. Federmeier. 2011. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology* 62. 621–647.

Kutas, Marta and Steven A. Hillyard. 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207(4427). 203–208.

Landau, Barbara, Linda B. Smith and Susan S. Jones. 1988. The importance of shape in early lexical learning. *Cognitive Development* 3(3). 299–321.

Lucy, John A. 1992. *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. New York, NY: Cambridge University Press.

Martín-Loeches, Manuel, Francisco Muñoz, Pilar Casado, Angela Melcón and Carlos Fernández-Frías. 2005. Are the anterior negativities to grammatical violations indexing working memory? *Psychophysiology* 42(5). 508–519.

Markman, Ellen M. and Jean E. Hutchinson. 1984. Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology* 16(1). 1–27.

Mizuguchi, Shinobu. 2004. *Individuation in numeral classifier languages: A case of Japanese classifiers and plurals.* Tokyo: Shohakusha.

Mueller, Jutta L., Anja Hahne, Yugo Fujii and Angela D. Friederici. 2005. Native and non-native speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience* 17(8). 1229–1244.

Münte, Thomas F. and Hans-Jochen Heinze. 1994. ERP negativities during syntactic processing of written words. In Hans-Jochen Heinze, Thomas F. Munte and George R. Mangun (eds.), *Cognitive electrophysiology*, 211–238. Boston, MA: Birkhauser.

Natsume, Soseki. 1906/1964. *Bottyan* [Master Daring]. Tokyo: Kaiseisha.

Neville, Helen, Janet L. Nicol, Andrew Barss, Kenneth I. Forster and Merrill F. Garrett. 1991. Syntactically based sentence processing classes: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 3(2). 151–165.

Osterhout, Lee and Phillip J. Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31(6). 785–806.

Osterhout, Lee and Linda A. Mobley. 1995. Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language* 34(6). 739–773.

Pelletier, Francis Jeffry. 1979. Non-singular reference. In Francis Jeffry Pelletier (eds.), *Mass terms: Some philosophical problems*, 1–14. Dordrecht: Kluwer Academic Publishers.

Quine, Willard Van Orman. 1960. *Word and object.* Cambridge, MA: MIT Press.

Quine, Willard Van Orman. 1969. *Ontological relativity and other essays*. New York: Columbia University Press.

Quine, Willard Van Orman. 1973. *Roots of reference*. La Salle, IL: Open Court.

Saalbach, Henrik and Mutsumi Imai. 2012. The relation between linguistic categories and cognition: The case of numeral classifiers. *Language and Cognitive Processes* 27. 381–428.

Sakai, Yumi, Kazuki Iwata, Jorge Riera, Xiaohong Wan, Satoru Yokoyama, Yoshiteru Shimoda, Ryuta Kawashima, Kei Yoshimoto and Masatoshi Koizumi. 2006. *Jishō kanren den-i de miru meishi to josūshi no shōgō purosesu – imi-teki shori ka bunpō-teki shori ka –* [An ERP study of the integration process between a noun and a numeral classifier: semantic or morpho-syntactic?] *Cognitive Studies* 13(3). 443–454.

Sakamoto, Tsutomu. 2015. Processing of syntactic and semantic information in the human brain: evidence from ERP studies in Japanese. In. Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: Mouton de Gruyter.

Samuelson, Larissa K. and Linda B. Smith. 1999. Early noun vocabularies: Do ontology, category organization and syntax correspond? *Cognition* 73(1). 1–33.

Sato, Kensuke and Etsuko Haryu. 2006. *Yoji ni okeru josūshi no rikai: sonzairon-teki kategori ni chūmokushite*. [Acquisition of numerical classifiers by Japanese preschoolers: Does ontological knowledge of the animate / inanimate distinction help them?] *Japanese Journal of Developmental Psychology* 17(3). 272–281.

Shipley, Elizabeth F. and Barbara Shepperson. 1990. Countable entities: Developmental changes. *Cognition* 34(2). 109–136.

Soja, Nancy N., Susan Carey and Elizabeth S. Spelke. 1991. Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition* 38(2). 179–211.

Subrahmanyam, Kaveri, Barbara Landau and Rochel Gelman. 1999. Shape, material, and syntax: Interacting forces in children's learning in novel words for objects and substances. *Language and Cognitive Processes* 14(3). 249–281.

Thierry, Guillaume, Dominique Cardebat and Jean-François Demonet. 2003. Electrophysiological comparison of grammatical processing and semantic processing of single spoken nouns. *Cognitive Brain Research* 17. 535–547.

Uchida, Nobuko and Mutsumi Imai. 1999. Heuristics in learning classifiers: The acquisition of the classifier system and its implications for the nature of lexical acquisition. *Japanese Psychological Research* 41(1). 50–69.

van Herten, Marieke, Herman H. J. Kolk and Dorothee J. Chwilla. 2005. An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research* 22(2). 241–255.

Yasunaga, Daichi and Tsutomu Sakamoto. 2007. On-line processing of floating quantifier constructions in Japanese: Using event-related brain potentials. *Journal of Japanese Linguistics* 23. 21–34.

Waxman, Sandra R. and Rochel Gelman. 1986. Preschoolers' use of superordinate relations in classification and language. *Cognitive Development* 1(2). 139–156.

Waxman, Sandra R. and Toby D. Kosowski. 1990. Nouns mark category relations: Toddlers' and preschoolers' word learning biases. *Child Development* 61(5). 1461–1473.

Yi, Byeong-uk. 2009. Classifiers and count nouns of Korean. *Proceedings of the 2009 summer conference of the International Society for Chomskyan Studies*, 125–140.

Yi, Byeong-uk. 2010. Numeral classifiers and the mass/count distinction. Ms. University of Toronto, 6 October.

Zhang, Niina Ning. 2012. De and the functional expansion of classifiers. *Language and Linguistics* 13(3). 569–582.

Zhang, Niina Ning. 2013. *Classifier structures in Mandarin Chinese*. Berlin: Mouton de Gruyter.

Shinji Fukuda, Suzy E. Fukuda and Tomohiko Ito

# 3 Grammatical deficits in Japanese children with specific language impairment

## 1 Introduction

Specific language impairment has been characterized as a congenital disorder of the normal course of language development in the absence of general cognitive disabilities, such as mental retardation, auditory impairment, autism, or any obvious neurological, psychological, or physical disorder that could account for the language deficit (Leonard 1998). In the literature, the terms 'developmental dysphasia', and 'language learning disability' have been widely used to roughly refer to the same condition. For clarity of presentation, 'specific language impairment' (henceforth 'SLI') is the term adopted in this chapter to describe this impairment.

It is widely believed that SLI is a disorder with a heterogeneous classification. Several researchers have indicated that SLI is a condition of abnormal language development affecting differing aspects of speech and language (Aram, Morris and Hall 1993; Rapin 1996; Conti-Ramsden, Crutchley and Botting 1997). The diagnosis of SLI is generally based on the fact that the language of the affected child develops late, and differs from normally-developing language, not on the linguistic properties of the SLI language itself. Therefore, a specific description of the disorder may not hold for all of the subtypes of SLI.

It is well documented that those with the most typical subtype of SLI have problems with some parts of grammatical development, especially with inflectional morphology. For instance, inconsistent use of inflectional morphemes such as the past tense *-ed*, the third person singular *-s*, and the plural *-s* is one of the most apparent problems reported in the literature (Leonard et al. 1992; Gopnik 1994; Rice, Wexler and Cleave 1995; Goad 1998; among others). Although children with SLI have problems with most inflectional affixes, the error rate varies among the inflectional affixes, as we shall later see in Section 3.1.

These problems and others have been accounted for by a diversity of etiological perspectives. In order to determine the underlying nature of SLI, primarily six linguistic accounts have been proposed: the Feature-blindness account, the Agreement Deficit account, the Structure-building Deficit account, the Extended Optional Infinitives account, the Implicit Rule Deficit account, and the Representational Deficit for Dependent Relations account. We will evaluate each of these accounts with Japanese SLI data. The language problems with SLI have sometimes been considered more of an epiphenomenon of a more general cognitive or peripheral processing problem. For example, they have been considered the result of a general information limita-

tion (Finneran, Francis, and Leonard 2009) or the result of an auditory processing deficit of rapid temporal speech sounds (Fellbaum, Miller, Curtiss and Tallal 1995). However, analyzing SLI data from a linguistic perspective will provide us with an alternative view of our language faculty. That is, it is possible to consider that the language problems observed in SLI are language-specific caused by a deficit to a particular part of the language module. Therefore, Japanese SLI data not only will contribute to the theory of specific language impairment in general, but also to a better understanding of just how the language module functions.

This chapter is organized as follows. In Section 2, we will provide an overview of the linguistically-principled data of Japanese-speaking children with SLI from previous studies as well as longitudinal data of a Japanese-speaking child with SLI. In Section 3, we will introduce six major linguistic accounts of SLI, and re-examine them in accordance with the Japanese SLI data provided in Section 2. Finally, we will provide concluding remarks in Section 4.

# 2 Japanese SLI data

Unlike the numerous studies on SLI in English, there have been a limited number of studies on SLI in Japanese. It is not the case, however, that there are only a small number of Japanese-speaking children with SLI. Perhaps, they have just been misdiagnosed because the concept of SLI is not well-recognized at either educational or clinical establishments in Japan.

In this section, we will first present an overview of Japanese SLI data from some major studies conducted over the past 15 years. We will then introduce some specific longitudinal data of a Japanese-speaking child with SLI.

## 2.1 Overview of Japanese SLI data

Fukuda and Fukuda (1999) was a preliminary study which investigated the linguistic characteristics of SLI in Japanese for the first time. In this pilot study, a battery of linguistically-principled tests was administered to eight Japanese-speaking children with SLI, ranging in age from 8;9 to 12;1, and eight age-matched children with normal language development. The battery was composed of tasks of syntactic comprehension, grammaticality judgment, Tense/Aspect production, grammaticality judgment of Tense/Aspect, among others.

The results of the syntactic comprehension task revealed that the children with SLI had difficulty comprehending certain utterances such as scrambled sentences and reversible passive sentences. The results of the grammaticality judgment task

revealed that the children with SLI had significant difficulty judging the ungrammaticality of certain sentences such as illicit passive and causative constructions, as well as illicit Case marker substitutions and omissions. The results of the Tense/Aspect sentence completion task showed that the children with SLI experienced great difficulty producing correct Tense and Aspect verb forms in the contexts that are exemplified in (1) and (2), respectively.

(1)

| | | | | |
|---|---|---|---|---|
| *Mainiti* | *Kazuo-kun wa* | *gakkoo* | *e* | *ik-u.* |
| every day | Kazuo-kun TOP | school | to | go-PRS |

'Every day, Kazuo goes to school.'

| | | | |
|---|---|---|---|
| *Kinoo mo* | *Kazuo-kun wa* | *gakkoo e* | ____. |
| yesterday-too | Kazuo-kun TOP | school to | |

'Yesterday too, Kazuo _____ to school.'
Target answer: *it-ta* 'went' (past tense form)

(2)

| | | | |
|---|---|---|---|
| *Mainiti* | *Kazuo-kun wa* | *onigiri o* | *tabe-ru.* |
| every day | Kazuo-kun TOP | riceball ACC | eat-PRS |

'Every day, Kazuo eats a riceball.'

| | | | | |
|---|---|---|---|---|
| *Ima mo* | *tyoodo* | *Kazuo-kun wa* | *onigiri o* | ____. |
| now too | right | Kazuo-kun TOP | riceball ACC | |

'Right now too, Kazuo _____ a riceball.'
Target answer: *tabe-te i-ru* 'be + eating' (present progressive form)

The most common type of error was the incorrect use of present (= non past) tense verb form in contexts where a past tense verb form or present progressive verb form was required. In the grammaticality judgment task of Tense/Aspect, the children with SLI also experienced difficulty judging the ungrammaticality of sentences in which a temporal adverb/adverbial phrase and a Tense/Aspect form of the predicate didn't match. In contrast, the age-matched children with normal language development performed all of the above tasks without any apparent difficulty. The mean percentage correct of the children with SLI was 68% while that of control children was 93%.

Fukuda and Fukuda (2001a) conducted an experimental study in order to further investigate the ability of children with SLI to form morphologically complex verbs. A sentence completion task was administered to six Japanese-speaking children with SLI, ranging in age from 7;4 to 12;1, and six age-matched children with normal language development (NLD). The children were asked to complete a sentence by supplying a missing suffix to the verb root/stem according to the picture which was shown to them. The stimulus sentences were presented to the child simultaneously both visually and orally. Some examples of the stimulus sentences are listed in (3).

(3) a. Intransitives
*Ki ga* *tao-*______.
tree NOM fall-______.
'The tree fell down.'
The target answer: *tao-re-ta* (fall-INTR[1]-PST)

b. Transitives
*Tanaka-san-tati ga* *ki o* *tao-*______.
Tanaka-san-PL NOM tree ACC fall-______.
'Lit.: Tanaka-san et al. fall$_{\mathrm{TRAN}}$ the tree.'
'Tanaka-san and his friend pulled the tree down.'
The target answer: *tao-si-ta* [fall-TR[1]-PST]

c. Passives
*Yamamoto-san ga* *Kazuko-san ni* *o(s)-*______.
Yamamoto-san NOM Kazuko-san DAT push-______.
'Yamamoto-san got pushed by Kazuko-san.'
The target answer: *os-are-ta* (push-PASS-PST)

d. Causatives
*Takako-san ga* *Emi-tyan ni* *guraundo o* *hasi(r)*______.
Takako-san NOM Emi-chan DAT track ACC run-______.
'Takako-san made Emi-chan run on the track.'
The target answer: *hasir-ase-ta* (run-CAUS-PST)

The results from this study illustrated that the children with SLI experienced significant difficulty forming lexicon-external complex verbs, namely passive verbs such as in (3c) and causative verbs such as in (3d) while they experienced much less difficulty forming lexicon-internal complex verbs, namely intransitive verbs such as in (3a) and transitive verbs such as in (3b).[2] The summary of the results is provided in Table 1.

---

**1** There are several intransitivizing suffixes (e.g., *-r-*, *-ar-*, *-re-*) and transitivizing suffixes (e.g., *-s-*, *-as-*, *-se-*) in Japanese. See Shibatani (1990) and Jacobsen (1992) for more details.

**2** Tanaka et al. (2001) conducted a syntactic production task with Japanese-speaking children with SLI, using the *Shitsugoshō Kōbun Kensa: Shōni-ban* ['The syntactic test of aphasia: child version'] (Fujita et al. 1984). In their study, seven 6-year old children with SLI also experienced difficulty with the production of passive and causative sentences, compared to intransitive and transitive sentences. The mean percentage correct of the children with SLI on the passive sentences was only 7.9% while that of age-matched control children was 52.4%. Their performance was significantly different. The mean percentage correct of the children with SLI on causative sentences was about 22% while that of age-matched control children was over 70%. However, their performance on causative sentences was not significantly different.

**Table 1:** Mean Percentage Correct

| | Intransitives | Transitives | Passives | Causatives |
|---|---|---|---|---|
| SLI | 90.0 | 77.4 | 32.7 | 42.3 |
| NLD | 95.0 | 87.6 | 91.6 | 94.0 |

The majority of errors were those in which the causative suffix -(*s*)*ase*- or the passive suffix -(*r*)*are*- was omitted where one of these suffixes was required. For example, many children with SLI produced *osi-ta* (push-PST) instead of *os-are-ta* (push-PASS-PST), (3c), and *hasit-ta* (run-PST) instead of *hasir-ase-ta* (run-CAUS-PST), (3d).[3]

Fukuda and Fukuda (2001b) conducted a follow-up study with eight Japanese-speaking children with SLI and eight age-matched children with normal language development, as well as with eight younger children with normal language development, and obtained basically the same results. The results from both of these studies suggest that the deficit of SLI affects the ability to construct implicit grammatical rules that are generated outside the domain of the lexicon, whereas their lexical operations for morphology that are generated within the domain of the lexicon appear to remain intact.

Fukuda, Fukuda, Ito and Yamaguchi (2007) conducted another experimental study in order to investigate whether or not Case was also affected in SLI. A sentence completion task was administered to three Japanese-speaking children with SLI, ranging in age from 9;7 to 13;3, and to five age-matched children with normal language development. Each child had to complete sentences by supplying missing Case markers according to the pictures that were presented to the child. The missing Case markers were one of the following three grammatical Case markers: the Nominative Case marker -*ga*, which attaches to the subject, the Accusative Case marker -*o*, which attaches to the direct object, and the Dative Case marker -*ni*, which attaches to the indirect object or to the subject in a tenseless subordinate clause. The stimulus sentences were presented to the child simultaneously both visually and orally. Some examples of the stimulus sentences are listed in (4) and (5). The examples in (4) are canonically word-ordered sentences in Japanese whereas those in (5) are scrambled sentences.

(4) a. *Hiro-tyan* (*ga*) *Aki-tyan* (*o*) *oikake-ta.*
Hiro-chan (NOM) Aki-chan (ACC) chase-PST
'Hiro-chan chased Aki-chan.'

3 By analyzing the spontaneous speech of two children with SLI, Otomo (2004) also found that they have problems with the production of the potential auxiliary verb -*re*(*ru*)/-*rare*(*ru*). However, their errors were inappropriate conjugations, not omissions of the potential auxiliary suffix. For example, they produced *torerareru* for *toreru* 'can take' and *syaberareru* for *syabereru* 'can speak'.

b. *Hiromi-san* (*ga*) *Satoru-san* (*ni*) *hako o* *hakob-ase-ta.*
Hiromi-san (NOM) Satoru-san (DAT) box ACC carry-CAUS-PST
'Hiromi-san made Satoru-san carry the box.'

c. *Kazuo-san* (*ga*) *Sizuka-san* (*ni*) *os-are-ta.*
Kazuo-san (NOM) Shizuka-san (DAT) push-PASS-PST
'Kazuo-san was pushed by Shizuka-san.'

(5) a. *Aki-tyan* (*o*) *Hiro-tyan* (*ga*) *oikake-ta.*
Aki-chan (ACC) Hiro-chan (NOM) chase-PST
'Hiro-chan chased Aki-chan.'

b. *Satoru-san* (*ni*) *Hiromi-san* (*ga*) *hako o* *hakob-ase-ta.*
Satoru-san (DAT) Hiromi-san(NOM) box ACC carry-CAUS-PST
'Hiromi-san made Satoru-san carry the box.'

c. *Sizuka-san* (*ni*) *Kazuo-san* (*ga*) *os-are-ta.*
Shizuka-san (DAT) Kazuo-san (NOM) push-PASS-PST
'Kazuo-san was pushed by Shizuka-san.'

The mean percentage correct of the children with SLI was 65%, while that of control children was 97%.[4] The results were further analyzed with respect to (i) simple sentences *vs.* complex sentences that consist of a main clause and a subordinate clause, and (ii) canonically word-ordered sentences *vs.* scrambled sentences. The most striking finding of this study was that the children with SLI had significant difficulty producing the correct Case markers in the reversible scrambled sentences, as exemplified in (5) above. More specifically, the mean percentage correct of the children with SLI for the reversible scrambled sentences was only 35% while that of the control children was 92%. The Case markers in the parentheses are those which children should have produced in the task. Regardless of the stimulus sentence type, simple or bi-clausal, the latter involving long distance scrambling, the children with SLI experienced significant difficulty producing the correct Case marker. Note that since the children were instructed to always provide some kind of Case marker after a noun, no omission errors were observed in this experiment. Very interestingly, two

**4** It appears that Japanese-speaking children with SLI also have problems with Inherent/Semantic Case, which is associated with semantic information. Murao, Matsumoto and Ito (2012) analyzed the spontaneous speech of two Japanese-speaking children with SLI, and found that they made errors with both structural Case (e.g., Nominative Case -*ga* and Accusative Case -*o*) and Inherent/Semantic Case (e.g., Locative Case -*de* 'at' and Conjunctive Case -*to* 'with'). However, their number of errors with structural Case was much greater than that with Inherent/Semantic Case. One 10-year old child made 37 errors with structural Case and only 8 errors with Inherent/Semantic Case, whereas the other 9-year old child made 103 errors with structural Case and 26 errors with Inherent/Semantic Case.

of the three children with SLI exhibited the same sort of error pattern. The typical errors with the scrambled sentences in (5), which were made by those two children, are provided in (6).

(6) a.

| | | |
|---|---|---|
| **Aki-tyan* (*ga*) | *Hiro-tyan* (*ni*) | *oikake-ta.* |
| Aki-chan (NOM) | Hiro-chan (DAT) | chase-PST |

‘Hiro-chan chased Aki-chan.’

b.

| | | | |
|---|---|---|---|
| **Satoru-san* (*ga*) | *Hiromi-san* (*ni*) | *hako o* | *hakob-ase-ta.* |
| Satoru-san (NOM) | Hiromi-san (DAT) | box ACC | carry-CAUS-PST |

‘Hiromi-san made Satoru-san carry the box.’

c.

| | | |
|---|---|---|
| **Sizuka-san* (*ga*) | *Kazuo-san* (*ni*) | *os-are-ta.* |
| Shizuka-san (NOM) | Kazuo-san (DAT) | push-PASS-PST |

‘Kazuo-san was pushed by Shizuka-san.’

It appeared as if when they were unsure about which Case marker to use they were using a rather unique strategy for Case marking. That is: (i) add the Nominative Case marker *-ga* to the first animate noun, and (ii) add the Dative Case marker *-ni* to the second animate noun in the linear word order.

Ito, Fukuda and Fukuda (2011) investigated whether or not Aspect was also affected in SLI. Utterances of spontaneous speech of two Japanese-speaking children with SLI were analyzed. One of the two children with SLI was a female junior high school student (15;6), and the other one was a male elementary school student (11;9). A sentence completion task was also conducted with these children with SLI, and the obtained results were compared with the results from 16 age-matched children with normal language development. The type of Aspect that was investigated in these studies was the V-*te* + *i-ru* form which denotes a continuance of an action or a state. In spontaneous speech, only a few errors with Aspect were observed. However, they experienced significant difficulty with the task, in which they were required to complete a sentence by supplying the missing aspectual suffix. The stimulus sentences were presented to the child visually. Incidentally, both children were not dyslexic. An example of a stimulus sentence is provided in (7).

(7)

| | | | | |
|---|---|---|---|---|
| (*Watasi wa*) | *kinoo kara* | *atarasi-i* | *kesigomu o* | *tuka*____. |
| (I TOP) | yesterday-since | new-PRS | eraser ACC | use____. |

‘Since yesterday, (I) use____ a new eraser.’

Target answer: *tukat-te i-ru* ‘have been using’ (continuous aspectual form)

The girl and the boy with SLI produced the correct response only 50% and 77.5% of the time, respectively, while the control children did so 95.9% of the time. The majority of errors were those with the past tense suffix *-ta* in contexts where the aspectual V-*te* + *i-ru* form was required.

## 2.2 Longitudinal data of a Japanese-speaking child with SLI

Ito, Fukuda and Fukuda (2009) investigated the linguistic aspects of a Japanese-speaking girl with SLI from the age of 9 to the age of 14. More specifically, the developmental changes in her performance with Tense, passives, Case, and demonstratives were examined, using sentence completion tasks, which required her to supply the missing element, and grammaticality judgment. Her lexical development was also examined using the Japanese version of the Picture Vocabulary Test – Revised (Ueno, Nagose, and Konuki, 2008)[5].

At the age of 9, the results of the sentence completion task revealed that the child experienced no difficulty producing a correct Tense form when frequent temporal adverbs such as *kinoo* 'yesterday' and *asita*, 'tomorrow' were used in the stimuli as exemplified in (8), but experienced significant difficulty when non-frequent temporal adverbial phrases such as *ima kara kokonoka mae ni* 'nine days before now' and *ima kara kokonoka ato ni* 'nine days after now' were used as exemplified below in (9). The stimulus sentences were presented to the child visually.

| | | | | |
|---|---|---|---|---|
| (8) | *Kinoo* | *Hanako wa* | *hon o* | *yo*(*m*)_____. |
| | yesterday | Hanako TOP | book ACC | read_____. |

'Yesterday, Hanako read a book.'
Target answer: *yon-da* / *yomi-masita* 'read' (past tense form)

| | | | | | | |
|---|---|---|---|---|---|---|
| (9) | a. | (*Watasi wa*) | *ima kara* | *kokonoka mae ni* | *ryokoo ni* | *i*(*k*)_____. |
| | | (I TOP) | now from | nine days before at | travel to | go_____. |

'Lit.: I went traveling nine days before now.'
'I traveled nine days ago.'
Target answer: *it-ta* / *iki-masita* 'went' (past tense form)

| | | | | |
|---|---|---|---|---|
| b. | (*Watasi wa*) | *ima kara kokonoka ato ni* | *ryokoo ni* | *i*(*k*)_____. |
| | (I TOP) | now from nine days after at | travel to | go_____. |

'Lit.: I will go traveling nine days after now.'
'I will travel nine days later.'
Target answer: *ik-u* / *iki-masu* 'will go' (non-past tense form)

In the grammaticality judgment task on Tense, she exhibited basically the same performance: experienced no difficulty when asked to judge the grammaticality of constructions with Tense when frequent temporal adverbs were used in the stimuli, but experienced significant difficulty when non-frequent temporal adverbial phrases were used.

5 The Picture Vocabulary Test – Revised (PVT-R) is a standardized language test to investigate the specific stage of a child's vocabulary development. In this test, the child is required to choose the correct picture among four, which corresponds to what s/he has heard.

When nonsense verbs and frequent adverbs were used in the stimuli, she also experienced significant difficulty on both the sentence completion and grammaticality judgment tasks of Tense. However, by the age of 14, her performance on both tasks significantly improved except on those with nonsense verbs, such as in (10). Actually, her performance with nonsense verbs got worse over the five year period. The percentages of correct responses were 50% at the age of 9, and 25% at the age of 14.

(10) *Taroo wa* *mainiti* *gakkoo e* *kim-u.*
Taro TOP every day school to kim-PRS
'Every day, Taro kims to school.'

*Kinoo mo* *Taroo wa* *gakkoo e* ____.
yesterday-too Taro TOP school to
'Yesterday too, Kazuo _____ to school.'
Target answer: *kin-da* (past tense form)

The child was also asked to complete a sentence by supplying the missing passive suffix to the verb stem based on the picture which was displayed simultaneously. The stimulus sentences were presented to the child visually. The results of this task revealed that she did relatively well when the word order was canonical from the age of 9 to 14, but at the age of 9 she experienced some difficulty when the word order of the passive sentences was reversed as in (5c), and did not improve at the age of 14. In the grammaticality judgment task on passive sentences, the child exhibited basically the same performance.

The results of the experiment on passives in this study appear to contradict the results of the sentence completion task with complex verbs in Fukuda and Fukuda (2001a, 2001b). However, there were two large differences between the two experiments. First, the number of stimuli, in which the target answer was a passive verb, was eight in the former experiment, whereas it was 30 in the latter. Therefore, in the former study some target passive verbs could possibly have been familiar ones which the child could have lexicalized as wholes. Secondly, the target answer was always a passive verb in the former experiment, whereas there was variety of different types of morphologically complex verbs in the latter (4 major types with some fillers). Therefore, it seems as if the child could have easily relied more on analogical knowledge in the former task to come up with the appropriate answer since all she had to do was to repeatedly produce a similar answer.

Table 2 shows the results of the child's development of demonstratives, namely *kono* 'this', *ano* 'that', *sono* 'its'[6], and *dono* 'which'.

6 Note that the usage of *sono* in Japanese is different from that of *its* in English.

**Table 2:** Results with demonstratives

| | 10;11 | 11;0 | 11;1 | 11;3 | 11;4 | 12;3 |
|---|---|---|---|---|---|---|
| Judgment | n./a. | 2/4 (50%) | 4/8 (50%) | n./a. | n./a. | 5/5 (100%) |
| Sentence Completion | 6/12 (50%) | n./a. | n./a. | 10/10 (100%) | 20/20 (100%) | 4/5 (80%) |

(Ito, Fukuda, and Fukuda 2009: 217)

In the judgment task, the child was asked whether or not the underlined demonstrative word in the sentence that corresponded with the picture was correct, and was also asked to provide an appropriate demonstrative phrase if she thought it was incorrect. An example of such a stimuli and its English translation is provided in Figure 1 and in (11), respectively.


**Figure 1:** A sample stimulus sentence of the judgment task on demonstratives

(11) **<u>Kono</u> hon wa omosiroi-yo.*
this book TOP interesting
'This book is interesting.'
Correct answer: *ano* 'that'

In the sentence completion task, the child was asked to complete the sentence in the picture by filling in the empty parentheses with one of the four demonstratives, namely *kono* 'this', *ano* 'that', *sono* 'its', and *dono* 'which'. An example of such a stimulus and its English translation is provided in Figure 2 and in (12), respectively.


**Figure 2:** A sample stimulus sentence of the sentence completion task on demonstratives

(12) The boy asks:
*Kono hon omosiroi?*
this book interesting
'Is this book interesting?'

The girl replies:
( ) *hon wa omosiroi-yo.*
book TOP interesting
'( ) book is interesting.'
Target answer: *sono* 'its'

As can be seen above in Table 2, her performance with demonstrative words improved rapidly in both judgment and sentence completion when the same stimuli were used.

Table 3 provides a comparison between the child's chronological age and lexical age. Her lexical age was measured by the Japanese version of the Picture Vocabulary Test – Revised. As can be seen in the results, her vocabulary developed quite rapidly, even more quickly than her development of demonstratives.

**Table 3:** Comparison of chronological age versus lexical age

| | | | |
|---|---|---|---|
| Chronological age | 9;00 | 9;10 | 10;10 |
| Lexical age | 5;10 | 6;07 | 10;05 |

(Ito, Fukuda, and Fukuda 2009: 217)

Fukuda, Fukuda, and Ito (2011) also examined the longitudinal data of the child with SLI in Ito, Fukuda, and Fukuda (2009). The study examined the comprehension of passive sentences. The child was asked to draw pictures depicting the contents of reversible passive sentences. The stimulus sentences were presented to the child visually. Some examples of the stimulus sentences are shown in (13). In the experiment itself, the Ns were actually the names of popular Japanese animated characters, such as *Nazonokusa* and *Rafushia*.

(13) a. *N1 ga N2 ni nage-rare-ru.*
N1 NOM N2 DAT throw-PASS-PRS
'N1 is thrown by N2.'

b. *N1 ga N2 ni hippa-rare-ru.*
N1 NOM N2 DAT pull-PASS-PRS
'N1 is pulled by N2.'

The experiment was conducted once a month for a period of five months from the time she was 10;1. The child's performance varied among the sentences from the first to the fourth time. However, on the fifth time, she came up with a specific


**Figure 3:** A picture drawn by a Japanese-speaking child with SLI (Fukuda, Fukuda and Ito 2011: 159)

compensatory strategy, and consequently was able to draw all pictures depicting the sentences correctly. More specifically, she explicitly said (in Japanese) "the second person in the sentences was the one who was actually doing something, whereas the first person was the one who was getting something done to him" (Fukuda, Fukuda and Ito 2011: 158). When passive sentences of reversed word order were presented, she used yet another compensatory strategy based on the previous one. More specifically, she connected the first person to the second person by drawing a reversible arrow between them, as exemplified in Figure 3.

The stimulus sentence, which corresponds with the above illustration, is provided in (14).

(14) *Yadon ni* *Ametama ga* *tukamae-rare-ru.*
Yadon DAT Ametama NOM catch-PASS-PRS
'Ametama is caught by Yadon.'

As can be seen in the above illustration, very interestingly, the child drew an arrow between the words, *Yadon* and *Ametama* (proper names of Japanese animated characters), in order to comprehend the scrambled passive sentence correctly. After that time, she correctly illustrated all the passive sentences.

## 3 The linguistic accounts

As previously mentioned in Section 1, the language impairment of SLI is often considered to be more of an epiphenomenon of a more general cognitive or peripheral

processing problem. The linguistic account, however, proposes that the deficit which results in SLI is language-specific caused by an impairment to a particular part of the language module. Therefore, it provides a very detailed account of the language impairment, aiming to explain the diverse errors characteristic of the language of children with SLI.

More precisely, linguistic accounts of SLI propose that the cause of SLI is an impairment in the language module that constrains the construction of grammars from the incoming linguistic data. It doesn't only assume that grammar is rule-governed. The linguistic account goes further and provides specific constraints on the content of these rules. It constructs a model specifying a hierarchy of constraints at different levels of the grammar predicting that one part of the grammar may be selectively impaired. All that operates within the linguistic module must be described in detail in order for this to be a valid account of this disorder. It must also be shown that the deficits of children with SLI can be accounted for in terms of these specific grammatical variables.

The impairment is postulated to be either an inability to construct a particular type of underlying abstract rule in the grammar or a delayed maturation of certain rules or categories of the grammar. Therefore, all instances of these kinds of rules are impaired independent of the surface form of the rule (Gopnik 1990b). Unfortunately, linguists disagree on the exact nature of this impaired underlying grammatical rule. It has been argued that it is syntactic-semantic feature marking (Gopnik 1990a, 1990b), that it is Agreement (Clahsen 1989, 1991), that it is structure-building (Guilfoyle, Allen, and Moss 1991; Rice 1992), that it is finiteness marking in matrix clauses (Rice, Wexler, and Cleave 1995; Rice and Wexler 1996; Rice, Wexler, and Hershberger 1998), that it is the ability to construct implicit grammatical rules (Gopnik 1994; Gopnik et al. 1997), and that it is the syntactic representation for grammatical dependent relations (van der Lely and Stollwerck 1997; van der Lely 1998; van der Lely and Battell 2003).

These six linguistic accounts were proposed based on SLI data in Indo-European languages, primarily English. They all provide detailed linguistic accounts, to differing degrees, of the actual language deficit of SLI in terms of language modularity. The data confirms that the linguistic accounts, postulating that a part (or some parts) of the underlying grammar is selectively impaired, can account for some of the errors characteristic of the disorder, but still do have some limitations. In this section, we will reexamine the validity of each account based on the Japanese SLI data presented in the previous section in addition to SLI data in Indo-European languages so that a universal account can be provided.

## 3.1 The Feature-blindness account

The Feature-blindness account was a hypothesis proposed by Gopnik (1990b) to describe the language of a single boy with SLI ('developmental dysphasia' in her

terms).[7] It is stated that, other than phonological information, at least three kinds of information must be provided in the lexicon: (i) grammatical class specifications, (ii) syntactic-semantic features, and (iii) specific semantic information. Gopnik predicted that the impaired grammatical characteristics typical of SLI were the result of a grammar without syntactic-semantic features, such as Tense, Aspect, Number, Person, Gender, among others in the lexicon. To avoid confusion, it should be noted that what she refers to as 'syntactic-semantic features' includes what are generally referred to as both 'morphosyntactic features' and 'grammatical features' in linguistic theory. She further claimed that because these syntactic-semantic features are absent, morphological rules that match features in the syntax were also not available. In contrast, there was no accompanying deficit in knowledge of the cognitive categories of the world because these categories are represented as part of the pure semantic specification of the word. In addition, the grammatical classes in the syntax, the thematic relations such as Agent and Theme in simple sentences, and the basic word order were all intact.

The Feature-blindness account well explained some of the manifestations of SLI since there clearly is a difference between the semantic holdings of children with SLI and their morphological production. For instance, they tend to express the notion of past by using temporal adverbs/adverbial phrases, instead of using the appropriate inflectional suffix *-ed* on the verb, as exemplified in (15).

(15) a. *Last time we arrive.*
b. *Last time I bring a one box of doughnuts.*
(Gopnik 1990b: 154)

Children with SLI also produce plural forms such as *trees* and *cops*, but do not reliably use the plural suffix *-s* to refer to more than one object. Children with SLI clearly understand the notion of plurality, and similar to the case of Tense marking, they tend to express it by using numeral quantifiers instead of using the appropriate inflectional suffix *-s* on the noun, as exemplified in (16).

(16) a. *I was make 140 box.*
b. *He only got two arena.*
(Gopnik 1990b: 147)

Therefore, feature marking theory confounded two phenomena that had to be distinguished from one another.

The advantage of the Feature-blindness account is that it can account for a wide diversity of errors which children with SLI exhibit. It can explain their errors with

7 See also Hegarty (2005) for a more or less similar account.

inflectional morphology such as Tense, Aspect, and Number marking. It also can explain their incorrect use of determiners as well as the lack of pronouns in their utterances.

This account, however, has problems accounting for the diverse error rates we find with different inflectional morphemes. For instance, it has been reported in the literature that children with SLI incorrectly omit the past tense *-ed* much more frequently than the progressive aspect marker *-ing* (Crystal 1987) while they also incorrectly omit the third person singular *-s* much more frequently than the plural *-s* (Rice and Oetting 1993).

With respect to the Japanese SLI children's data, the Feature-blindness account appears to be able to account for the difficulties with Tense and Aspect marking as well as with complex predicate formation. However, it is not clear whether or not this account can explain their syntactic problems such as their difficulty with the comprehension of reversible passive sentences and the production of Case markers in reversible scrambled sentences as were presented in (13a) and (5c), respectively, and which have been repeated below.

(13) a. *N1 ga* *N2 ni* *nage-rare-ru.*
N1 NOM N2 DAT throw-PASS-PRS
'N1 is thrown by N2.'

(5) c. *Sizuka-san* (*ni*) *Kazuo-san* (*ga*) *os-are-ta.*
Shizuka-san (DAT) Kazuo-san (NOM) push-PASS-PST
'Kazuo-san was pushed by Shizuka-san.'

Gopnik herself later reports that English-speaking children with SLI also have problems with syntactic comprehension tasks. For example, they experienced difficulty comprehending reversible passive sentences such as *The boy is pushed by the girl* (Gopnik 1999). It may be possible to provide an adequate explanation for such performance if we assume that NP-movement such as scrambling and passivization is also a feature-driven operation. However, since Gopnik herself did not provide an analysis for the syntactic comprehension problems of children with SLI in her work, we are also unable to provide an examination.

## 3.2 The Agreement Deficit account

Clahsen (1989) examined grammatical errors produced by German-speaking children with SLI, and argued that dysphasic children (he refers to 'SLI' as 'developmental dysphasia') have problems in establishing grammatical Agreement relations. His interpretation of Agreement is much larger than the general definition of Agreement in linguistic theory. That is "structural relations between two elements in which one

element asymmetrically controls the other" (Clahsen 1989: 916). He predicted a lack of Agreement between Number and Gender with nouns and their corresponding adjectives and articles in the noun phrase, namely that between the Case-marked noun and the verb.

To investigate these predictions, Clahsen (1991) analyzed two sets of data: spontaneous speech samples from 10 German-speaking children with SLI and spontaneous speech samples and elicitation data from 20 children with SLI studied longitudinally over a period of one year. He examined properties of syntax and inflectional morphology such as word order, constituent structure, negation, question formation, Case marking, verb morphology, and plural morphology.

His results indeed supported his account: Gender and Number Agreement in the noun phrase were often incorrect, and subject-verb Agreement caused great difficulty. The children used full noun phrases and pronouns appropriately in head-final position as required in German. However, within the noun phrases, they had problems with determiners. More specifically, they frequently omitted articles in obligatory contexts in various contexts. In addition, they had great problems with the use of correct Gender and Number markings. Some examples of their incorrect use of Gender marking are provided in (17).

(17) a. *und* <u>*de*</u> *tild*
'and the sign' = (We need) the sign

b. *un* <u>*das*</u> *po letzt*
'and the bum hurt' = her bottom is hurt.

c. *ich* <u>*die*</u> *Lehrer bin*
'I the teacher am'
(Clahsen 1991: 134)

With respect to the structure of NPs, they rarely produced complex NPs such as Det+Adj+N.

Concerning verbal elements, the children with SLI used simple verbs, prefixed verbs, and modal verbs. In contrast, very few cases of copulas and auxiliaries were found. Some examples of omissions of copular and auxiliary verbs in obligatory contexts among their utterances are shown in (18) and (19), respectively.

(18) *hase lieb*
'hare sweet' = The hare is sweet.
(Clahsen 1991: 140)

(19) *schinken aufgessen*
'ham eated' = (The dog) has eaten all of the ham.
(Clahsen 1991: 140)

The proportion of deleted verbal elements gradually decreased over time.

Clahsen's data showed that the children with SLI also had numerous problems with the use of Case markings required in German. For example, in contexts requiring Accusative Case for the object, they often used Case-neutral markers as in (20a) or Dative Case markers as in (20b).

(20) a. *der mann noch mal rausnehmen*
= (We) take out the man.

b. *ich dir hinführen*
= I am leading you there.
(Clahsen 1991: 157)

In contexts requiring Dative Case for the object, they also often used Case-neutral markers as in (21a), or Accusative Case markers as in (21b).

(21) a. *du besser helf ich*
= I help you better.

b. *du mis ein geb*
= You give me one.
(Clahsen 1991: 158)

The children with SLI most often used a "binary Case system" with Nominative Case for the subject and either Accusative or Dative Case for the object. However, they sometimes used the Accusative Case marker or Dative Case marker for the subject, as exemplified in (22a) and (22b), respectively.

(22) a. *uns auch so was*
= We've got something like that too.

b. *ihm kipsbein nachher kommt*
= He'll get his leg plastered afterwards.
(Clahsen 1991: 158)

In addition, there were very few instances of Case Agreement on elements such as the determiner or the adjective within the NP. In fact, none of the children with SLI in his study had successfully acquired the Case Agreement paradigm within the NP.

It was subject-verb Agreement that caused the most problems for the children with SLI. In German, the verb form needs to agree with the grammatical person and the number of the subject. There are five suffixes that mark subject-verb Agreement, namely *-ø*, *-e* (schwa), *-st*, *-t*, and *-n*. The use of the suffixes *-ø*, *-e*, and *-n* was most frequently observed whereas the use of suffix *-st*, was almost never seen. The percentages correct were very low, except with *-t*. With the exception of one child named

*Petra*, all of the children with SLI did not show any improvement on these verbal inflections over time.

Lastly, the children with SLI showed evidence of difficulty with word order, placing the verb in the final position SOV, and not the second position SVO. Clahsen concluded that children with SLI had problems mainly in the areas of inflectional morphology and with function words. Therefore, he claimed that the focus of the deficit in SLI was clearly in grammatical Agreement.

The data of German-speaking children with SLI in Clahsen's studies appears to support the Agreement Deficit account. However, his account cannot account for much of the English SLI data. For example, as previously stated, one of the most typical errors, which English-speaking children with SLI make, is Tense marking. They often produce a bare stem form in the past context, as exemplified by *My dad wash his car yesterday*. The form of the verb *wash* needs to be matched with the temporal adverb *yesterday*, but it is not the case that the temporal adverb determines the form of the verb. The past tense form is required because the event expressed by the sentence happened in the past. Therefore, Tense marking errors by English-speaking children with SLI cannot be explained by the Agreement Deficit account. In addition, the same problems arise for Aspect marking and Number marking.

Furthermore, there has been a report of some evidence which further refutes the Agreement Deficit account. Rice and Oetting (1993) studied spontaneous language samples of 81 children with SLI, who had less problems with Agreement within the noun phrase (agr) such as *two cup-s*, but had great problems with Agreement between the noun and the verb (Agr) such as *she run-s*. Their results demonstrate that Agreement is not a unitary phenomenon across the grammar. The Agreement Deficit account also incorrectly predicts that children with SLI would experience great difficulty matching the type of subject and its corresponding copula, *be*, and consequently would produce errors like *I is* . . . , *You am* . . . , *Mary are* . . . , and so on. As far as we are aware, no examples of this sort of error have been reported in the literature.

The Agreement Deficit account cannot account for some of the Japanese SLI data, either. This account certainly cannot explain their inability to produce appropriate lexicon-external complex predicates, in which the attachment of the passive or causative suffix is required, as was illustrated in (3c) and (3d) and has been repeated below, since complex word formation has no property related to Agreement.

(3) c. Passives

| | | |
|---|---|---|
| *Yamamoto-san ga* | *Kazuko-san ni* | *o(s)-______.* |
| Yamamoto-san NOM | Kazuko-san DAT | push-______. |

'Yamamoto-san got pushed by Kazuko-san.'

The target answer: *os-are-ta* (push-PASS-PST)

(3) d. Causatives

| | | | |
|---|---|---|---|
| *Takako-san ga* | *Emi-tyan ni* | *guraundo o* | *hasi*(*r*)_____. |
| Takako-san NOM | Emi-chan DAT | track ACC | run-______. |

‘Takako-san made Emi-chan run on the track.’
The target answer: *hasir-<u>ase-ta</u>* (run-CAUS-PST)

In addition, the Agreement deficit account cannot explain the asymmetric performance on Case marking by the Japanese-speaking children with SLI. Recall that they performed relatively well with the production of the correct Case marker in canonically word-ordered sentences such as (4a) whereas they had significant difficulty with the production of the correct Case marker in reversible scrambled sentences such as (5a), which have been repeated below.

(4) a.

| | | |
|---|---|---|
| *Hiro-tyan* (*ga*) | *Aki-tyan* (*o*) | *oikake-ta.* |
| Hiro-chan (NOM) | Aki-chan (ACC) | chase-PST |

‘Hiro-chan chased Aki-chan.’

(5) a.

| | | |
|---|---|---|
| *Aki-tyan* (*o*) | *Hiro-tyan* (*ga*) | *oikake-ta.* |
| Aki-chan (ACC) | Hiro-chan (NOM) | chase-PST |

‘Hiro-chan chased Aki-chan.’

Similarly, this account fails to account for the atypical development in the comprehension of reversible passive sentences such as *Nazonokusa ga Rafushia ni nage-rare-ru.* (‘Nazonokusa is thrown by Rafushia.’), which the child with SLI exhibited. It appears that she had problems in assigning thematic roles in passive sentences, in which there had been NP-movement (passivization), that is not related to Agreement.

## 3.3 The Structure-building Deficit account

In linguistic theory, lexical items are divided into two syntactic categories, namely lexical categories and functional categories. Lexical categories include nouns, verbs, adjectives, and prepositions. These categories contain a great amount of semantic information and little grammatical information. Functional categories include inflections (Infl),[8] determiners (Det), complementizers (Comp), and Case. In contrast to lexical categories, these categories contain a great amount of grammatical information and little or no semantic information. Some kinds of verb movements are linked to the development of functional categories. Consequently, in German, word order is argued to be SOV until Infl develops, and only after the development of Infl can

**8** Note that inflections (Infl) have been replaced by Tense (T) in contemporary linguistic theory.

word order be expected to change to SVO. The development of the Nominative Case on the subject is also argued to follow the development of Infl.

The Structure-building account makes several assumptions about language acquisition. Functional categories are hypothesized to emerge later than lexical categories in the course of normal language development (Radford 1990; Guilfoyle and Noonan 1992; Vainikka 1993/94). The emergence of functional categories is determined largely by a maturational schedule. Although there are individual differences among children, only lexical categories emerge around the age of 20 months, and functional categories begin emerging around the age of 24 months in English (Radford 1990). Saliency in the language of input plays a role in determining the timing of the appearance of functional categories in different languages. Therefore, for example, in languages where the functional categories occur syllabically or carry more meaning for the language, such categories would be expected to develop a little earlier. In German, for instance, where Case marking is more salient than in English, it does in fact develop earlier.

The Structure-building Deficit account assumes that SLI is a result of the reduced ability to build syntactic structure. As a result, utterances of children with SLI lack functional categories since these categories are located in higher positions of syntactic hierarchical structure, and the syntactic derivational process takes place bottom-to-top. In other words, this account predicts that the nature of SLI is the delayed maturation of functional categories (Guilfoyle, Allen, and Moss 1991; Rice 1992[9]). The lexical stage, for example, is argued to be prolonged and the grammar to resemble telegraphic speech with a lack of inflection such as Tense, Agreement, Aspect, and Number, as well as with a lack of function words such as determiners. However, children with SLI have the ability to construct basic syntactic structures because their ability to assign thematic roles such as Agent and Theme in simple sentences remains intact. Variability is expected among children with SLI with the grammar having 'fossilized' at some point in the acquisition sequence. One would not expect to see the development of the complementizer without the development of past tense marking.

Rice (1992) tested this account by examining 81 spontaneous speech samples from English-speaking preschool children with SLI. She compared these with 92 spontaneous speech samples from language-matched children with normal language development. Her results supported the predictions of the account. There was great variability within her pool of 81 children with SLI. Some children showed difficulty with the use of determiners and Agreement marking on nouns, namely with Number marking on nouns such as *these cup-s*. They took the form of bare stems with omitted affixes, whereas some other children showed no problems at all. Subject-

**9** Rice later revised her account, and argued that children with SLI have specific problems with Spec-head relations within a functional phrase (i.e., a phrase headed by a functional category). See Rice (1994) for more details.

verb Agreement, however, was a problem for all the children with SLI. For instance, the Agreement suffix *-es* in parentheses in (23) was omitted in their speech. Verbs were hardly ever marked for inflection. Primarily, the children used bare stem forms.

(23) *My brother wash(-es) his car everyday.*

The assignment of thematic roles posed no particular problems for either group, as expected. Consequently, Rice concluded that the functional category model allowed us to observe some interesting morphosyntactic asymmetries in their performance.

Difficulty with Agreement relations, which children with SLI exhibit, may be regarded as having more to do with the delayed maturation of certain functional categories than with the complete lack of their presence. An advantage of the Structure-building Deficit account, she claims is "that it identifies particular questions and at the same time places them in a much broader picture than that of localized, individual morphemes" (Rice 1992).

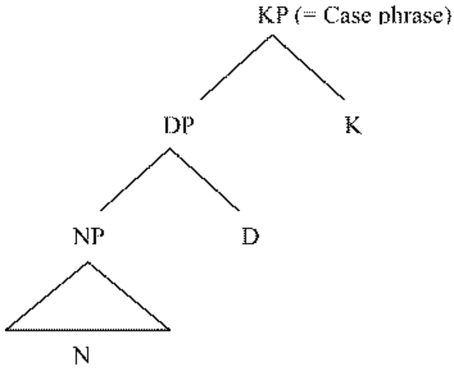
In contrast to the Agreement Deficit account, the Structure-building Deficit account seems to be able to account for a much broader range of the problems of English-speaking children with SLI, such as problems with past tense marking, third person singular Agreement, plural marking, and determiners (See Nakayama and Yoshimura's chapter in this volume on similar errors in L2 English by Japanese EFL learners).

However, there are some problems with the Structural-building Deficit account. First, this account cannot explain the optionality of inflectional errors. If certain functional categories have not yet emerged, such functional categories should rarely if ever appear in children's utterances. Gopnik's data clearly demonstrate variability. Her morphosyntactic investigations reveal that all inflected forms are present, but are not used consistently (Gopnik 1992). Such performance implies the existence of functional categories, i.e., a mature adult grammar. In fact, the studies conducted in Leonard (1995) and Eyer and Leonard (1995) demonstrated that all functional categories were present in the speech of children with SLI, but were less frequently used in obligatory contexts, compared with language-matched younger children with normal language development. Note that the maturational hypothesis was primarily proposed to account for development that occurs in children with normal language development at very young ages. Furthermore, there is German data that shows both for children with SLI and for those with normal language development that verb movement can be found to occur before the development of the inflectional category, Infl (Clahsen 1991). This observation is the exact opposite of what the maturational hypothesis predicts.

Regarding the Japanese SLI data, at first glance, it appears that this account adequately explains the errors with Case marking which the children with SLI exhibit, such as in (5). This is because the Structure-building Deficit account predicts that the higher the syntactic position of a functional category is the more likely it is

to be omitted, and that Case is generated under the highest head position, i.e., KP is higher than DP. This is shown in (24), according to Travis and Lamontagne (1992).

(24)




*Modified for head-final languages.

However, if the functional category, Case (K), is absent in the grammar of children with SLI, Case markers should be always missing in utterances of children with SLI. As we have seen in (4) and (5), that is not the case. Recall that the children with SLI performed relatively well producing correct Case markers in canonically word-ordered sentences such as in (4) whereas they had significant difficulty doing so in reversible scrambled sentences such as in (5).

More importantly, the Structure-building Deficit account cannot explain their inability to produce appropriate lexicon-external complex predicates, in which the attachment of the passive or causative suffix is required such as in *os-are-ta* (push-PASS-PST) and *hasir-ase-ta* (run-CAUS-PST), because the passive and causative suffixes are both auxiliary verbs which are considered to be a lexical category, not a functional category.

## 3.4 The Extended Optional Infinitives account

According to Wexler (1994), at a young age, children go through a period of time when they often use infinitival verb forms in matrix clauses where a finite form is required. He named this period the Optional Infinitive (henceforth OI) stage of language development. He assumes that children in the OI stage do not yet realize that it is obligatory to mark finiteness such as Tense and Agreement in matrix clauses. He argues that children with normal language development go through this stage, but successfully acquire the correct use of finite forms around 5 years of age.

Rice, Wexler, and Cleave (1995), Rice and Wexler (1996), Rice, Wexler, and Hershberger (1998), and Rice, Wexler, and Redmond (1999) argue that children with SLI go

through an Extended Optional Infinitive (henceforth EOI) stage in which the period of the OI stage is extended. What they claim with the EOI account is that children with SLI remain in an OI stage for a longer period of time, compared to children with normal language development. As a result, children with SLI often produce root forms in matrix clauses. In other words, English-speaking children with SLI sometimes produce uninflected verbs without a required suffix such as the past tense *-ed* and the third person singular *-s* in obligatory contexts because they are in the EOI stage, as exemplified in (25), respectively.

(25) a. *My mom cook*(*-ed*) *dinner last night.*
b. *My dad walk*(*-s*) *a mile every morning.*

They further argue that the omissions of the auxiliaries *do*/*be* and the copula *be* in obligatory contexts are also a result of their inability to realize that finiteness marking is obligatory in matrix clauses. In other words, they argue that children with SLI often make grammatical errors like the examples in (26) because the children are in the EOI stage.

(26) a. *Tom* (*did*) *not go to school yesterday.*
b. *Mary* (*is*) *eating cookies.*
c. *I* (*am*) *always very happy.*

The words in parentheses in (26) are the elements which children with SLI often omit. They conclude that the use of incorrect of Tense/Agreement markers and the omissions of the auxiliaries *do*/*be* and the copula *be* in finite clause by children with SLI is no different from the use of such forms by younger children with normal language development.

The EOI account is, perhaps, the least plausible account of SLI. The most outstanding problem with this proposal is the very small range of language difficulties it accounts for. It can only explain the incorrect use of root forms of the verb in matrix clauses where inflected forms are required. It may also be able to explain the omissions of the auxiliaries *do*/*be* and the copula *be* in matrix clauses. However, as has been noted, the language difficulties children with SLI experience are much broader. As we have observed in previous sections, English-speaking children with SLI also experience difficulty with plural marking (Leonard et al. 1992; Goad 1998), as well as numeral quantifier-noun Agreement (Rice and Oetting 1993). In addition, German-speaking children with SLI also experience difficulty with determiner-noun Agreement and Gender Agreement in noun phrases (Clahsen 1989; 1991). Case marking has also been reported to be problematic for English-speaking children with SLI (Radford 2005) and German-speaking children with SLI (Clahsen 1989; 1991). Furthermore, it has been widely reported that English-speaking children with

SLI also have problems with Voice, more specifically with the comprehension of reversible passive sentences (van der Lely 1994, 1996; Gopnik 1999). Apparently, the EOI account cannot account for any of these impaired grammatical characteristics of SLI.

With respect to the Japanese SLI data, the same sort of problem arises. That is the fact that the EOI account can only explain very few problems which children with SLI exhibit. This account may be able to explain their problems with verb morphology such as Tense marking and Aspect marking. It should be noted that since the root in Japanese is a bound morpheme it never surfaces independently. When they make errors, therefore, they use a verb with an inappropriate suffix. In contrast, the EOI account cannot explain their other difficulties such as problems with the formation of lexicon-external complex predicates such as *os-<u>are-ta</u>* (push-PASS-PST) and *hasir-<u>ase-ta</u>* (run-CAUS-PST), the production of correct Case markers in scrambled sentences, as was illustrated in (5), and the comprehension of reversible passive sentences such as *Nazonokusa ga Rafushia ni nage-rare-ru* ('Nazonokusa is thrown by Rafushia.') in which NP-movement (i.e., passivization) is involved.

There is another serious problem with the EOI account. If this account is correct, sooner or later, children with SLI should eventually grow out of the OI stage. However, this doesn't seem to be the case. It has been reported that language difficulties of SLI persist, at least, for decades (Stothard, et al. 1998; Johnson, et al. 1999; Ito, Fukuda, and Fukuda 2009), if not throughout their entire lives (Gopnik 1990a; Gopnik and Crago 1991).

## 3.5 The Implicit Rule Deficit account

The Implicit Rule Deficit account was originally proposed based on a theory of learnability of inflected words called the dual mechanism hypothesis (Pinker and Prince 1988; Pinker 1991, 1997). The dual mechanism model incorporates both a computational component, which contains abstract symbolic rules and representations, and an associative memory system with certain properties of connectionist models. The claim is that regular (e.g., *wash*/*wash-ed*) and phonologically similar irregular (e.g., *break*/*broke*) inflected words are processed differently. More specifically, regular verbs are derived through the application of a procedural rule, add *-ed* to the verb stem, and phonologically similar irregular verbs derived through analogical learning devices within an associative memory. In addition, phonologically unrelated irregular verbs (e.g., *go*/*went*) are claimed to be derived through pure memorization within rote memory.

In the Implicit Rule Deficit account, Gopnik (1992, 1994, 1996) argues that individuals with SLI are unable to reliably formulate implicit grammatical rules for certain properties such as Tense and Number. She hypothesizes that individuals with SLI can learn individual words such as *walked* and *books* as unanalyzed wholes

by means of this association network stored in declarative memory, but cannot generalize from these individual instances to construct procedural symbolic rules that would operate on an abstract category, for example, a rule for constructing regular past tense (e.g., *wash* + *-ed* → *wash-ed*).

In order to provide a linguistically-principled account of the underlying grammar of the individuals with SLI, Gopnik (1992, 1994), and Gopnik and Crago (1991) examined a wide range of both spoken and written production data as well comprehension data. It was collected over a period of two and a half years from thirty members of a three-generation family, sixteen of whom had been diagnosed with SLI: the family known as the 'KE family'. It consisted of administered tests, such as grammaticality judgment tasks, grammaticality rating tasks, auditory comprehension tests and various production tasks as well as spontaneous speech samples.

The results from several different tests converge to support the hypothesis that for both verbs and nouns, these individuals cannot construct implicit rules that govern morphological Agreement. They did show evidence of being able to learn some of these inflected words by memorizing them as unanalyzed single lexical items, and constructing association networks stored in declarative memory. In grammaticality judgment tasks of morphosyntactic features, their performance was no better than chance. Their ability to correct errors on morphosyntactic features was significantly poorer than that of the individuals with normal language development. In an auditory comprehension task of singular *vs.* plural, there was no significant difference between the responses of the individuals with SLI and those with normal language development. In a Wug Test[10], administered to test the hypothesis that the individuals with SLI lexicalize *s*-marked forms, and do not generate them from a pluralization rule, there again was a significant difference between the individuals with SLI and those with normal language development (Gopnik 1992).

The most striking difference in performance between the individuals with SLI and those with normal language development was in their rating of the stem form for both regular and irregular verbs. The individuals with SLI, unlike those with normal language development, did not judge that a stem form in a temporally past sentence was ungrammatical. They did not appear to have dichotomous ratings for verb forms. In their spontaneous speech, as well, they often produced a non-past form in a temporally past context, as exemplified in (27). In contrast, they never produced a past tense form in a present context.

**10** The original Wug Test was reported in Berko Gleason (1958). It was a sentence completion task, which was designed to investigate the formation of the plural and the use of other inflectional morphemes in English-speaking children. In this task, the child is presented with a picture of a non-existent creature, and told by an experimenter 'This is a wug'. Another drawing, which contains two of the non-existent creatures, is then shown to the child, and the experimenter says 'Now there are two of them', 'There are two ____'. Children, who have already acquired the plural suffix *-s*, are expected to respond 'wug-s'.

(27) *She remembered when she hurts herself the other day.*
*Then Goldilocks sit down.*
*Then we went canoeing really fast and then I fall in three times.*
(Gopnik 1994: 123)

In addition, in their unconstrained narrative speech, they also often produced a non-past form in a temporally past context, as exemplified in (28).

(28) *The boy climb up the tree and frightened the bird away.*
*They call the ambulance and the ambulance came.*
*He did it then he fall.*
*The neighbors phoned an ambulance because the man fall off the tree.*
(Gopnik 1994: 126)

Data from correction tests showed that they are unsure of the form that a verb should have. Gopnik (1992, 1994) argues that these results clearly demonstrate that the individuals with SLI cannot reliably 'manipulate Tense marking' on verbs to produce sentences that are grammaticality correct with respect to Tense.

Gopnik argues all of this verb data is consistent with a model in which Tense marking is not obligatory in the grammar of the individuals with SLI. However, there is evidence that they somehow have acquired the knowledge of the correct form for past tense verbs since they sometimes use them correctly in their spontaneous speech. An analysis of longitudinal writing data revealed that they were learning the past tense forms of regular verbs one at a time (Gopnik 1992).

Gopnik (1992, 1994) concludes that this linguistically-principled analysis of the data demonstrates that individuals with SLI lack the ability to construct implicit symbolic rules in their grammar. Only by using declarative memory, and by constructing association networks, can individuals with SLI learn some properties of language.

The Implicit Rule Deficit account can explain the inconsistent use of inflected forms, which English-speaking children with SLI exhibit, such as Tense, Aspect, Agreement, and Number marking. In addition, this account can provide an adequate explanation for the difference in error rates between regular and irregular past-tense verbs. As previously mentioned, there have been some reports which show that individuals with SLI perform relatively better with irregular verbs than with regular verbs with regard to Tense marking (Gopnik 1994: Ullman and Gopnik 1999). If individuals with SLI indeed memorize past tense forms of regular verbs as unanalyzed wholes by means of this association network stored in declarative memory, as Gopnik argues, we should find a strong frequency effect for both regular and irregular past tense forms, which was not found in individuals with normal language development. That is exactly what we find (Ullman and Gopnik 1999; van der Lely and Ullman 2001).

Nevertheless, similar to the Feature-blindness account, it is not clear how to explain the diverse error rate we find with different inflectional morphemes with this account. As previously stated, children with SLI incorrectly omit the past tense *-ed* and the third person singular *-s* much more frequently than the progressive aspect marker *-ing* and the plural *-s*. One possible answer would be that children with SLI produce inflected forms by using explicit knowledge in the similar manner that second language learners learn grammatical rules of foreign languages in formal language classes (Paradis and Gopnik 1997). For example, they can produce past tense forms using explicit knowledge like "Add an *-ed* to verbs in which the event happened in the past", or "Add an *-s* to nouns in which there are more than two items." In order to use such explicit knowledge, semantic notion plays an important role. This would explain their poor performance with the third person singular *-s* since it is purely a syntactic marker, which has absolutely zero semantic information. The notion of Tense seems more abstract than that of Number because the latter is visually recognizable while the former is not. This might explain why children with SLI perform better with Number marking than with Tense marking. However, what remains unsolved with such an explanation is the difference in performance between Tense and Aspect marking. The semantic values of Tense and Aspect are both relatively abstract, but as previously noted, children with SLI experience greater difficulty with the former when compared to the latter.

The Implicit Rule Deficit account seems to be able to explain the incorrect use of function words such as determiners (e.g., *a*, *an*, and *the*) and the omission of the auxiliary verb *be* in progressive contexts since Gopnik herself does not explicitly state that the implicit grammatical rules which children with SLI are unable to reliably formulate are limited to morphosyntactic operations.

Turning to the Japanese SLI data, the Implicit Rule Deficit account seems to be able to explain the difficulties with Tense marking as well as with Aspect marking. Recall that the child with SLI in Ito, Fukuda, and Fukuda (2009), however, performed well on Tense production when the temporal adverbs in the stimuli were frequent such as *kinoo* 'yesterday' as in (8), which is recalled below.

| | | | | |
|---|---|---|---|---|
| (8) | *Kinoo* | *Hanako wa* | *hon o* | *yo*(*m*)_____. |
| | yesterday | Hanako TOP | book ACC | read_____. |
| | 'Yesterday, Hanako read a book.' | | | |
| | Target answer: *yon-da* / *yomi-masita* (past tense form) | | | |

These results contradict those from the Tense production experimental data in Fukuda and Fukuda (1999), and also appear to be in contradiction with the severe problems experienced by English-speaking children with SLI with Tense marking that have been well documented in the literature. At the moment, it is not clear whether or not the good performance of the child is specific to this particular child,

or is due to the fact that, unlike in English, the bare stem verb never surfaces without a Tense marker in Japanese. The results from the experimental study on complex predicate formation demonstrate that children with SLI are not only unable to reliably formulate implicit grammatical rules of inflectional morphology but also those of derivational morphology such as passive and causative suffixation such as in (3c) and (3d), respectively.

As the examination in this subsection illustrates, overall, the Implicit Rule Deficit account can provide an adequate explanation for the morphosyntactic problems of children with SLI. One might wonder, however, whether or not this account can also explain the syntactic problems of Japanese-speaking children with SLI such as the problems with the comprehension of reversible passive sentences such as in (13) and the production of Case markers in reversible scrambled sentences such as in (6). As previously mentioned, English children with SLI also experience difficulty with the comprehension of reversible passive sentences. Recall that Gopnik theorizes that children with SLI are unable to reliably formulate implicit grammatical rules, so that her theory may be also applicable to syntactic rules. Nevertheless, since she herself has not provided a detailed explanation about how her account could explain the syntactic problems of children with SLI, we are unable to examine her claims, and must leave them to future research.

## 3.6 The Representational Deficit for Dependent Relations account

The Representational Deficit for Dependent Relations (henceforth RDDR) account proposes that the deficit behind at least a subtype of SLI is in the syntactic computational system (van der Lely 1998).[11] The subtype of SLI that she refers to is the so-called Grammatical SLI (G-SLI), whose prominent problems are persistent grammatical difficulties in both the production and comprehension of language at the levels of morphosyntax and syntax. More specifically, the RDDR account argues that children with G-SLI have a modular language deficit with syntactic dependent structural relationships between constituents. By adopting the Minimalist framework (Chomsky 1995), the RDDR account further argues that the grammatical difficulties of children with G-SLI are primarily due to an optional movement operation (i.e., Move) in syntactic derivation.

Over the years, van der Lely and her colleagues investigated the grammatical competence of children with G-SLI using a wide variety of language tasks. The

---

**11** van der Lely and her colleagues later revised this account, and argued that the grammatical difficulties, which children with grammatical SLI experience, are due to a deficit in representing linguistic structural complexity in the three components of the computational grammatical system, namely syntax, morphology, and phonology. See Marshall and van der Lely (2007), Theodoros and van der Lely (2007), Marshall, Theodoros, and van der Lely (2007), and van der Lely, Jones, and Marshall (2011) for more details.

children with G-SLI exhibited a significant delay in grammatical development. In addition to inflectional morphology (e.g., the 3rd person, singular *-s*, and the past tense *-ed*) which are the most typical problems in children with SLI, children with G-SLI have great difficulty forming syntactically complex structures involving embedded phrases such as (29a) when the PP is embedded in the NP (van der Lely 1998), Binding Principles (Chomsky 1981) such as identifying the antecedent of anaphoric reflexives (e.g., *himself/herself*) and pronouns (e.g., *him/her*) (van der Lely 1998; van der Lely and Stollwerck 1997), comprehension of reversible passive sentences such as (29b) (van der Lely 1994, 1996), and production of object *wh*-questions such as (29c) (van der Lely 1998; van der Lely and Battell 2003).

(29) a. [$_{\text{NP}}$ *The cat* [$_{\text{PP}}$ *with the blue blanket*]] *is jumping on the bed.*
b. *The boy is pushed by the girl.*
c. *Who did Mrs. Peacock see in the lounge?*

The RDDR account argues that the impairment of inflectional morphology is a result of optional head-to-head movement. To be more precise, the impairment of Tense is due to optional V-to-T movement while the impairment of Agreement is due to optional V-to-Agr movement. It also elegantly explains the syntactic deficits which children with G-SLI exhibit. For example, the problematic comprehension of reversible passive sentences can be accounted for by optional NP-movement to A-position (i.e., Spec of TP) whereas the problematic production of object *wh*-questions can be accounted for by optional movement of the *wh*-operator to the A-bar-position (i.e., Spec of CP) and I-to-C movement of *do*.

The optional syntactic movement proposed in the RDDR account is able to explain the optionality of language performance which children with G-SLI do exhibit. With respect to inflectional morphology, it is well-known that it is not the case that children with SLI always omit inflectional affixes. They do produce correct inflected forms, but sometimes omit inflectional affixes, resulting in ungrammatical sentences. In other words, they cannot consistently use correct inflected forms. With respect to the comprehension of reversible passive sentences and the production of object *wh*-questions, it is not the case that they never understand reversible passive sentences nor never produce object *wh*-questions. They sometimes understand reversible passive sentences, and sometimes produce object *wh*-questions. However, their level of competence is always much lower than that of age-matched children with normal language development. Such inconsistent performance of children with G-SLI can be accounted for by optional syntactic movement.

The RDDR account, however, appears to have problems explaining some of the manifestations of children with SLI. It is well documented in the literature that children with SLI incorrectly omit most inflectional affixes in obligatory contexts, but omit different inflectional affixes at different degrees. As repeatedly mentioned

in the previous sections, English-speaking children with SLI omit the past tense *-ed* and the third person singular *-s* much more frequently than the progressive aspect marker *-ing* and the plural *-s*. If the RDDR account is correct, this means that optional head-to-head movement takes place at different degrees. It is not clear at all what would trigger such a different frequency of syntactic movement. It would also be hard to explain the difference in error rates between regular and irregular past-tense verbs with optional head-to-head movement. In addition, the inconsistent use of independent function words such as the determiner (e.g., *a*, *an*, and *the*) and the omission of the auxiliary verb *do*/*be* and the copula *be* would remain unaccounted for with optional head-to-head movement.

Furthermore, it is argued that the lack of complex NPs such as an NP with an embedded PP (e.g., *The cat with the blue blanket*) in the utterances of children with G-SLI is due to the fact that they can only build simple structures which involve basic local dependencies, and not more complex long dependencies (van der Lely and Stollwerck 1997; van der Lely 1998). It should be noted, however, although this explanation is indeed a syntactic representational problem, no syntactic movement is involved.

As previously stated, children with G-SLI have problems with Binding Principles. For example, in the sentence (30) children with G-SLI sometimes accepted *Mowgli* as well as *Baloo Bear* as the antecedent of the anaphoric reflexive *himself* which is a violation of Principle A in Binding Theory[12] (van der Lely and Stollwerck 1997; van der Lely 1998).

(30) *Mowgli says Baloo Bear is tickling himself.*

van der Lely and Stollwerck (1997) argue that children with G-SLI are unable to compute complex syntactic structures. As a result, they have problems with complex syntactic dependencies (i.e., long-distance dependencies). It is not clear to us, however, how the violation of Principle A in Binding Theory is caused by a complex syntactic dependency since the relationship between the anaphoric reflexive (e.g., *himself*/*herself*) and its antecedent is rather local.

The RDDR account also appears to be able to explain the Japanese SLI data. Recall that the Japanese children with SLI had problems with the comprehension of reversible passive sentences such as in (13) and the production of Case markers in reversible scrambled sentences in (6). The RDDR account can explain these problems because both passive sentences and scrambled sentences involve NP-movement. This account can also explain their problems with Tense marking and Aspect marking because they both involve head-to-head movement. In addition, this account can explain their problems with complex verb formation in which attachment of the passive or causative suffix to the verb is required because complex verb formation

**12** Binding Principle A: An anaphor is bound in its governing category (Chomsky, 1981).

involves V-to-V head movement. The RDDR account, however, has some problems explaining the Japanese data. Fukuda and Fukuda (2001a, 2001b) report that the children with SLI showed relatively good performance forming intransitive and transitive verbs such as in (3a) and (3b), respectively which are also morphologically complex in Japanese. This account fails to account for these results because these verbs are also formed by head-to-head movement (i.e., V-to-*v* movement in a split VP configuration) according to Nishiyama (1998) and Hasegawa (1999).

# 4 Conclusion

There have been numerous linguistically-principled studies of SLI in English (Gopnik 1994; 1999; Rice, Wexler, and Cleave 1995; Ullman, and Gopnik 1999; van der Lely 1998; 2005; among many others). There has also been a large number of linguistically-principled studies of SLI in highly inflected languages such as German (Clahsen 1989; 1991), French (Jakubowicz 2003), Greek (Dalalakis 1994), and Dutch (Kenneth, Schaeffer, and Gerard 2004). Nevertheless, there have only been a small number of studies on the linguistic characteristics of SLI in agglutinative languages such as Turkish, Korean, and Japanese, and in polysynthetic languages such as Mohawk, Wichita, and Kiowa. It would be quite interesting to see how SLI manifests itself in these languages because SLI data from these languages may provide us with valuable information which is simply unattainable in English and highly inflected languages. For example, in Japanese, we can test how children with SLI perform with Case marking, various types of verbal morphological processes, and scrambled sentences, to name just a few.

In this chapter, we reexamined the validity of six linguistic accounts of SLI with Japanese SLI data in addition to SLI data in other languages, primarily that of English. None of the linguistic accounts claimed to be able to perfectly predict and substantiate the full range of impaired grammatical characteristics of children with SLI. Some more than others, however, appear to have succeeded in doing so. Among the six linguistic accounts, it turned out that the Implicit Rule Deficit account appears to be the most adequate hypothesis of SLI. Among the six linguistic accounts, it could best account for the impaired grammatical characteristics of SLI. It could also provide an explanation for the wide diversity of grammatical errors which children with SLI exhibit. Regarding inflectional morphology, it could explain the optionality of their errors (= the inconsistent use of inflected forms) and the different error rates between regular and irregular past verb forms. In addition, it also could account for the difficulty they experience with derivational morphology as well as with independent words such as determiners and pronouns. Furthermore, it appeared able to explain their syntactic problems with the addition of a slight modification to the account. Lastly, and most importantly, this account was able to provide an adequate explanation of the Japanese SLI data as well.

As can be seen, a comprehensive examination of SLI in Japanese has provided us with an opportunity to both better evaluate various hypotheses of the deficit proposed with data from other languages, and to formulate a more universal linguistically-principled account for SLI phenomena. This in turn will provide us with a deeper understanding of how language is processed in the human brain. As very few studies are available on Japanese SLI, further linguistically-principled investigations from a variety of perspectives are definitely needed in future research.

## Acknowledgments

We are very grateful to Mineharu Nakayama for his insightful comments. We wish to thank all of the participants of the Japanese studies cited in this chapter, and the speech-language pathologists/instructors for both their collaboration and use of their facilities for testing. We also would like to thank Ayumi Kimura, Kana Kobari, Yoshimi Kannou, Tomoko Aita, and Yuko Yamaguchi for their assistance in collecting the data in the studies. Many of our research projects cited in this chapter were supported by the Grant-in-Aid for Scientific Research (C) from the Japan Society for the Promotion of Science, #1520266, #16530420, #19530584, and #22530696, for which we are truly grateful.

## References

Aram, Dorothy M., Robin Morris and Nancy Hall. 1993. Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech and Hearing Research* 36. 580–591.

Berko Gleason, Jean. 1958. The child's learning of English morphology. *Word* 14. 150–177.

Bortolini, Umberta, Maria Cristina Caselli and Laurence B. Leonard. 1997. Grammatical Deficits in Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research* 40. 809–820.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Clahsen, Harald. 1989. The grammatical characterization of developmental dysphasia. *Linguistics* 27. 897–920.

Clahsen, Harald. 1991. *Child language and developmental dysphasia. Linguistic studies of the acquisition of German*. Amsterdam: John Benjamins Publishing.

Conti-Ramsden, Gina, Alison Crutchley and Nicola Botting. 1997. The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research* 40. 765–777.

Crago, Martha and Shanley E. M. Allen. 1994. Morphemes gone askew: Linguistic impairment in Inuktitut. *McGill Working Papers in Linguistics* 10. 206–215.

Crystal, David. 1987. *Clinical linguistics*. Baltimore, MD: Arnold.

Dalalakis, Jenny. 1994. Familial language impairment in Greek. *McGill Working Papers in Linguistics* 10. 216–228.

Eyer, Julia A. and Laurence B. Leonard. 1995. Functional categories and specific language impairment: A case study. *Language Acquisition* 4. 177–203.
Fellbaum, Christiane, Steven Miller, Suzan Curtiss and Paula Tallal. 1995. An auditory processing deficit as a possible source of SLI. In Dawn MacLaughlin and Susan McEwen (eds.), *Proceedings of the 19th annual Boston University Conference on Language Development*, 204–215. Cascadilla Press, Somerville, MA.
Finneran, Denise. A., Alexander L. Francis, and Laurence B. Leonard. 2009. Sustained attention in children with specific language impairment. *Journal of Speech, Language, and Hearing Research* 52. 915–929.
Fujita, Ikuyo, Takako Miyake, Yukinobu Nakanishi, Etsuko Imamura, Mariko Morita, Noriko Sato, Hiroko Hayata and Keiko Kodama. 1984. *Shitsugoshō Kōbun Kensa: Shōni-ban* [‘The syntactic test of aphasia: child version’]. Nihon Chōnō Gengo-shi Kyōkai: Shitsugoshō Kensa-hō Iinkai.
Fukuda, Suzy E. and Shinji Fukuda. 1999. Specific language impairment in Japanese: A linguistic investigation. *NUCB Journal of Language, Culture, and Communication* 1. 1–25.
Fukuda, Shinji and Suzy E. Fukuda. 2001a. The acquisition of complex predicates in Japanese specifically language-impaired and normally developing children. *Brain and Language* 77. 305–320.
Fukuda, Shinji and Suzy E. Fukuda. 2001b. An asymmetric impairment in Japanese complex verbs in specific language impairment. *Cognitive Studies* 8. 63–84.
Fukuda, Shinji, Suzy E. Fukuda, Tomohiko Ito, and Yuko Yamaguchi. 2007. Grammatical impairment of Case in Japanese children with specific language impairment (In Japanese). *The Japan Journal of Logopedics and Phoniatrics* 48. 95–104.
Fukuda, Suzy E., Shinji Fukuda and Tomohiko Ito. 2011. Atypical development of the passive construction in a Japanese child with specific language impairment. *Ars Linguistica* 18. 152–163.
Goad, Heather. 1998. Plurals in SLI: Prosodic deficit or morphological deficit? *Language Acquisition* 7. 247–284.
Gopnik, Myrna. 1990a. Feature-blind grammar and dysphasia. *Nature* 344. 715.
Gopnik, Myrna. 1990b. Feature blindness: A case study. *Language Acquisition* 1. 139–164.
Gopnik, Myrna. 1992. *Linguistic properties of genetic language impairment*. Paper presented at Conference of American Association for the Advancement of Science. Chicago, IL, February.
Gopnik, Myrna. 1994. Impairment of tense in a familial language disorder. *Journal of Neurolinguistics* 8. 109–133.
Gopnik, Myrna. 1999. Familial language impairment: More English evidence. *Folia Phoniatrica et Logopaedica* 51. 5–19.
Gopnik, Myrna and Martha Crago. 1991. Familial aggregation of a developmental language. *Cognition* 39. 1–50.
Gopnik, Myrna, Jenny Dalalakis, Suzy E. Fukuda, and Shinji Fukuda. 1997. The biological basis of language: Familial language impairment. In Myrna Gopnik (ed.), *The inheritance and innateness of grammars*, 111–140. New York, NY: Oxford University Press.
Guilfoyle, Eithne, Shanley Allen and Siobhan Moss. 1991. Specific language impairment and the maturation of functional categories. Paper presented at the 16th Annual Boston University Conference on Language Development. Boston, MA.
Guilfoyle, Eithne and Máire Noonan. 1992. Functional categories and language acquisition. *Canadian Journal of Linguistics* 37. 241–272.
Hasegawa, Nobuko. 1999. *Seisei nihongo-gaku nyūmon* [‘Introduction to generative Japanese linguistics’]. Tokyo: Taishūkan.
Hegarty, Michael. 2005. *A feature-based syntax of functional categories: The structure, acquisition and specific impairment of functional systems*. Berlin: Mouton de Gruyter.

Ito, Tomohiko, Shinji Fukuda and Suzy E. Fukuda. 2009. Differences between grammatical and lexical development in Japanese specific language impairment: A case study. Poznań Studies in Contemporary Linguistics 45. 211–221.

Ito, Tomohiko, Shinji Fukuda and Suzy E. Fukuda. 2011. Aspect in Japanese children with SLI. *Asia Pacific Journal of Speech, Language, and Hearing* 14. 23–29.

Jakubowicz, Celia. 2003. Computational complexity and the acquisition of functional categories by French-speaking children with SLI. *Linguistics* 41. 175–211.

Jacobsen, Wesley M. 1992. *The transitive structure of events in Japanese*, Tokyo: Kuroshio Publishers.

Johnson, Carla, Joseph H. Beitchman, Arlene Young, Michael Escobar, Leslie Atkinson, Beth Wilson, E. B. Brownlie, Lori Douglas, Nathan Taback and Isabel Lam. 1999. Fourteen-year follow-up of children with and without speech/language impairments. *Journal of Speech, Language and Hearing Research* 42. 744–760.

Leonard, Laurence B. 1998. *Children with specific language impairment*. Cambridge, MA: MIT Press.

Leonard, Laurence B. 1995. Functional categories in the grammars of children with specific language impairment. *Journal of Speech and Hearing Research* 38. 1270–1283.

Leonard, Laurence B., Umberta Bortolini, M. Cristina Caselli, Karla K. McGregor, and Letizia Sabbadini. 1992. Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition* 2. 151–179.

Marshall, Chloe R., Marinis Theodoros and Heather K. J. van der Lely. 2007. Passive verb morphology: The effect of phonotactics on passive comprehension in typically developing and grammatical-SLI children. *Lingua* 117. 1434–1447.

Marshall, Chloe R. and Heather K. J. van der Lely. 2007. The impact of phonological complexity on past tense inflection in children with grammatical-SLI. *Advances in Speech-Language Pathology* 9. 191–203.

Murao, Aimi, Sachiyo Matsumoto-Shimamori and Tomohiko Ito. 2012. Case-marker errors in the utterances of two children with specific language impairment (In Japanese). *The Japan Journal of Logopedics and Phoniatrics* 53. 194–198.

Nakayama, Mineharu and Noriko Yoshimura. 2015. The modularity of grammar in L2 acquisition. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Nishiyama, Kunio. 1998. *The morphosyntax and morphophonology of Japanese predicates.* Ithaca: Cornell University dissertation.

Otomo, Kiyoshi. 2004. Syntactic characteristics in specific language impairment in Japanese: Analysis of the speech of twins with difficulties in syntax (In Japanese). *Tōkyō Gakugei Daigaku Tokushu Kyōiku Kenkyū Shisetsu Kenkyū Hōkoku* 3. 1–9.

Paradis, Michel and Myrna Gopnik. 1997. Compensatory strategies in genetic dysphasia: declarative memory. *Journal of Neurolinguistics* 10. 173–186.

Pinker, Steven. 1991. Rules of language. *Science* 253. 530–535.

Pinker, Steven. 1997. Words and rules in the human brain. *Nature* 387. 547–548.

Pinker, Steven and Alan Prince. 1988. On language and connectionism: Analysis of the parallel distributed processing model of language acquisition. *Cognition* 28. 73–193.

Radford, Andrew. 1990. *Syntactic theory and acquisition of English syntax*. Oxford: Blackwell Publishers.

Radford, Andrew. 2005. Accusative subjects and defective clauses in the grammar of an English-speaking child with specific language impairment. In Masahiko Minami, Harumi Kobayashi, Mineharu Nakayama, and Hidetoshi Shirai (eds.), *Studies in Language Sciences* (4), 3–41. Tokyo: Kuroshio Publishers.

Rapin, Isabelle. 1996. Practitioner Review: Developmental language disorders: A clinical update. *Journal of Child Psychology and Psychiatry* 37. 643–655.

Rice, Mabel L. 1992. *Grammatical categories of specifically language-impaired children*. Paper presented at the Bruton Conference, Dallas, TX.
Rice, Mabel L. 1994. Grammatical categories of children with specific language impairments. In Ruth V. Watkins and Mabel L. Rice (eds.), *Specific language impairments in children: Current directions in research and intervention*, 69–89. Baltimore, MD: Paul H. Brookes.
Rice, Mabel L. and Janna B. Oetting. 1993. Morphological deficits in children with SLI: Evaluation of number marking and agreement. *Journal of Speech and Hearing Research* 36. 1249–1257.
Rice, Mabel L. and Kenneth Wexler. 1996. A phenotype of specific language impairment: Extended optional infinitives. In Mabel L. Rice (ed.), *Toward a genetics of language*, 215–237. Mahwah, NJ: Lawrence Erlbaum Associates.
Rice, Mabel L., Kenneth Wexler and Patricia Cleave. 1995. Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research* 38. 850–861.
Rice, Mabel L., Kenneth Wexler and Scott Hershberger. 1998. Tense over time: The longitudinal course of tense acquisition in children with specific language impairments. *Journal of Speech, Language, and Hearing Research* 41. 1412–1431.
Rice, Mabel L., Kenneth Wexler and Sean M. Redmond. 1999. Grammaticality judgments of an extended optional infinitive grammar: Evidence from English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research* 42. 943–961.
Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.
Stothard, Susan E., Margaret J. Snowling, Dorothy V. M. Bishop, Barry B. Chipchase and Carole A. Kaplan. 1998. Language impaired preschoolers: A follow-up into adolescence. Journal of Speech, Language, and Hearing Research 41. 407–418.
Tanaka Welty, Yumiko, Jun Watanabe, Kazuko Shirakuni and Lise Menn. 2001. Japanese preschool children with specific language impairment: A preliminary exploration of language characteristics (In Japanese). *The Japanese Journal of Communication Disorders* 18. 2–9.
Theodoros, Marinis and Heather K. J. van der Lely. 2007. On-line processing of *wh*-questions in children with G-SLI and typically developing children. *International Journal of Language and Communication Disorders* 42. 557–582.
Travis, Lisa and Greg Lamontagne. 1992. The Case filter and licensing of empty K. *Canadian Journal of Linguistics* 37. 157–174.
Ueno, Kazuhiko, Naoko Nagose and Satoru Konuki. 2008. *PVT-R Kaiga goi hattatsu kensa* [*PVT-R: Picture Vocabulary Test – Revised* (Japanese version)]. Tokyo: Nihon Bunka Kagakusha.
Ullman, Michael T. and Myrna Gopnik. 1999. Inflectional morphology in a family with inherited specific language impairment. *Applied Psycholinguistics* 20. 51–117.
van der Lely, Heather K. J. 1994. Canonical linking rules: Forward versus reverse linking in normally developing and specifically language impaired children. *Cognition* 51. 29–72.
van der Lely, Heather K. J. 1996. Specifically language impaired and normally developing children: Verbal passive vs. adjectival passive sentence interpretation. *Lingua* 98. 243–272.
van der Lely, Heather J. K. 1998. SLI in children: movement, economy, and deficits in the computational- syntactic system. *Language Acquisition* 7. 161–192.
van der Lely, Heather K. J. and Jackie Battell. 2003. *Wh*-movement in children with grammatical SLI: a test of RDDR Hypothesis. *Language* 79. 153–181.
van der Lely, Heather K. J., and Linda Stollwerck. 1997. Binding theory and specifically language impaired children. *Cognition* 62. 245–290.
van der Lely, Heather K. J. and Michael T. Ullman. 2001. Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes* 16. 177–217.
van der Lely, Heather K. J. 2005. Grammatical-SLI and the computational grammatical complexity hypothesis. *Révue Fréquences* 17. 13–20.

van der Lely, Heather K. J., Melanie Jones, and Chloe R. Marshall. 2011. Who did Buzz see someone? Grammatically judgement of *wh*-questions in typically developing children and children with Grammatical-SLI. *Lingua* 121. 408–422.

Vainikka, Anne. 1993/94. Case in the development of English syntax. *Language Acquisition* 3. 257–325.

Wexler, Kenneth, Jeannette Schaeffer and Gerard Bol. 2004. Verbal syntax and morphology in typically developing Dutch children and children with SLI: How developmental data can play important role in morphological theory. *Syntax* 7. 148–198.

Keiko Murasugi

# 4 Root infinitive analogues in Child Japanese

## 1 Introduction

Root Infinitives (RIs) are non-finite (infinitival) verb forms used in matrix (root) clauses, i.e., a context where they cannot appear in adult grammar, by children around two years of age. Root Infinitives are attested in very young children's speech across a wide variety of languages. Although the use of non-finite verbs in root contexts by very young children is a universal phenomenon, there are morphological variations associated with the different verbal systems in the children's target languages. RIs can be infinitives, bare verbs, participles, or certain (surrogate) "finite" forms.

In some languages with relatively rich morphology such as Dutch (Haegeman 1995; Blom and Wijnen 2000) and French (Krämer 1993; Rasetti 2003), among others, children may optionally use the infinitival forms of inflection on the verb, rather than finite ones.[1]

(1) #*Peter* *bal* *pakken.* (2;1) (Dutch)
Peter ball get-INF
'Peter (wants to) get the ball.'
(Blom and Wijnen 2000)

(2) #*Dormir* *petit* *bébé.* (1;11) (French)
sleep-INF little baby
'A little baby sleeps.'
(Guasti 2004)

On the other hand, in languages which are relatively poor in inflectional morphology like English, non-finite verbs appear in the finite (root) contexts as bare verbs. In adult English, infinitive forms are generally the bare stems, and English-speaking children produce the bare stems within the age range of 20–36 months as shown in (3).

1 Abbreviations used in the glosses are as follows: ACC = Accusative Case, ASP = Aspect morpheme, DAT = Dative Case, INF = Infinitive, MIM = Mimetic word, MOOD = Mood marker, NEG = Negation, NOM = Nominative Case, PRS = Present, PST = Past, REQ = Request, SFP = Sentence final particle.

(3) a. #*Eve sit floor.* (1;7) (English)
(Brown 1973)

b. #*That truck fall down.* (2;0) (English)
(Sano and Hyams 1994)

Just like in English, very young children speaking Swahili also omit functional elements such as tense and subject agreement (Deen 2002).[2] An equivalent non-finite stage has also been identified for children acquiring languages that do not have an infinitive construction. In Modern Greek, for example, a bare subjunctive/perfective is reported to be the Root Infinitive analogue (RI analogue) (Varlokosta, Vainikka and Rohrbacher 1996; Hyams 2002).

There are also many languages whose RI analogue is the "full" form. Kim and Phillips (1998) suggest that the RI analogue for Korean is the verb stem with the mood marker *-e*. Bar-Shalom and Snyder (2001) report that children speaking Russian produce two forms of RIs: infinitives in a root clause and imperative forms. Salustri and Hyams (2003) also observe that the proportion of imperatives is significantly higher than that of RIs. According to Salustri and Hyams (2003, 2006), Italian-speaking children begin using imperatives before age two, and the verbs have appropriate morphology.

(4) *dammi!* (1;10)
give-to $\text{me}_{\text{cl}}$
'Give it to me.'
(Salustri and Hyams 2003)

Similarly, Lillo-Martin and Quadros (2009) and Chien (2008) propose that imperative forms are RI analogues in sign languages (American Sign Language (ASL)/Brazilian Sign Language (LSB)) and Chinese, respectively. Grinstead (1998), Bel

2 Deen (2002) typologically classifies child languages into three types: languages that allow "true" RIs such as German and French, languages that have no RI phenomenon such as Italian and Japanese, and languages like Swahili whose very early non-finite verb forms appear with bare verbs. In this paper, we assume that not only the children speaking Italian, Japanese, Spanish, Catalan, but also children speaking such pro-drop languages as Chinese, ASL, and Turkish, for example, go through the RI analogue stage. (See Murasugi, Fuji and Hashimoto 2010; Murasugi, Nakatani and Fuji 2012; among others.)

**Table I:** Typology of Root Infinitives (Deen 2002)

| True RI Languages | | Non-RI Languages | | Bare Verb Languages | |
|---|---|---|---|---|---|
| Dutch | French | Catalan | Italian | English | Inuktitut |
| German | Icelandic | Japanese | Spanish | Quechua | Sesotho |
| Russian | Swedish | | Siswati | Swahili | |

(2001), and Montrul (2003) find that imperatives are quite frequent in the early stage and decrease over time in Spanish and Catalan.

Dutch has been considered to be a typical RI language, but Wijnen, Kempen, and Gillis (2001) report that verbal forms resembling imperatives are found, in addition to the Infinitive forms, at the early two-word stage. If that is the case, then Dutch-speaking children produce the imperative forms as well as the infinitive forms as their first verbs.

(5) “. . . Starting with the early two-word stage, forms resembling imperatives were discarded from the analyses, as it is unclear whether they are finite or non-finite.”
(Wijnen, Kempen and Gillis 2001)

The findings independently obtained from Russian, Italian, and Dutch described above should not be labeled coincidental. The very early non-finite verbs do not necessarily appear in a single form per language, and the “(apparently) fully conjugated” forms seem to be chosen as the RI analogue in more than just a few languages.

It is well known that there are some salient morpho-syntactic and semantic properties of RIs, as listed in (6).

(6) a. RIs are tenseless verbs in root contexts.
b. At the RI stage, no T-related/C-related items are found.
c. RIs are produced to describe events in real time, that is, as an on-going activity in past, present, or future that the child is involved in (Aspect Effects).
d. RIs occur in modal contexts (the Modal Reference Effects (MREs)).
e. RIs are restricted to event-denoting predicates (the Eventivity Constraint).
f. Head Merger is not available during the RI analogue stage.

For RIs, two peculiar types of contextual interpretation have been identified. One type refers to so-called extensional contexts, whereby RI analogues are produced to describe events in real time, that is, as an on-going activity in past, present, or future that the child is involved in. The other type of interpretation refers to so-called intentional contexts, whereby RI analogues are produced as a result of children’s intention, desire, or volition, in various irrealis modal contexts. This is termed the Modal Reference Effects (MREs) (Hoekstra and Hyams 1998).

The MREs, described in (6d), mean that RIs typically have a modal or irrealis meaning, expressing volition or request (Hoekstra and Hyams 1998; among others). Observe the example in (7) from Dutch.

(7) #*Vrachtwagen emmer doen.* (2;4) (Dutch)
truck bucket do-INF
Context: Matthijs (speaker) wants the investigator to put the truck in the bucket.
(Blom and Wijnen 2000)

Besides the MREs, it has been also widely observed that RIs are largely restricted to eventive predicates as shown in (6), whereas finite verbs can either be eventive or stative. Early eventive verbs tend to receive a modal meaning with overwhelming frequency, and this is termed the Eventivity Constraint (Hoekstra and Hyams 1998). As is clear from the English case given in (3), the head merger between V and T (ense) is not fully available either during the stage of RI analogues (Phillips 1995, 1996; Murasugi and Fuji 2009).

It has been also pointed out that RIs do not occur in interrogative sentences with *wh* nor with T-related elements such as *be*-copula and auxiliaries. According to Haegeman (1995), *wh*-questions are rarely produced by children at two to three years of age.

(8) #*Wie staat daar?* (2;6) (Dutch)
who stands there?
'Who stands there?'
(Haegeman 1995)

When *wh*-questions are produced by young children, the main verbs used in the *wh*-questions are finite, as shown in (8) and Table 1. This is termed Crisma's effect.

**Table 1:** Finiteness in declaratives and questions:
Dutch (Haegeman 1995, modified in Phillips 1995, 1996)

| Hein 2;4–3;1 | *+finite* | *–finite* | %-finite |
|---|---|---|---|
| All clauses | 3768 | 721 | 16% |
| *wh*-questions | 88 | 2 | 2% |

Total = 4579, $\chi^2 = 12.71$, $p < 0.001$

Infinitive verbs are cross-linguistically common in child language, and the phenomenon is widespread. RIs have some salient morpho-syntactic and semantic properties. The stage ends fairly consistently by age of three or so. Obviously, the phenomenon exhibits some deficiency in the functional structure of children who use RIs, but what exactly does it mean that the RIs are not marked for tense or agreement?

## 2 Problems

The Root Infinitive (RI) phenomenon has occupied a central place in the generative studies of language acquisition (Rizzi 1993/1994; Wexler 1994; Hoeskstra and Hyams 1998; among others). Nevertheless, there are several mysterious aspects of RIs that

have not received adequate descriptions and explanations. For instance, RIs have been considered to be optional phenomena, because children speaking English, for example, use both non-finite and finite verbs at the stage of RIs. However, it is not crystal clear what "optionality" exactly means.

Second, the cross-linguistic distribution of RIs is gradient (Guasti 2002): in the acquisition of non-pro-drop languages, e.g., English, Dutch, and German, children have quite a long period of RI use, sometimes extending over three years. In contrast, in such pro-drop languages as Japanese, Korean, Italian, Catalan, and Spanish, there is a very short RI analogue stage.

Furthermore, in such pro-drop languages as Japanese and Korean, the RI analogue stage starts very early and ends before age two. Grinstead (2000), for example, finds that Spanish and Catalan-speaking children at a very early stage lack contrastive use of tense and number morphology, but this stage ends around 1;10. This raises a question: How is the property of "pro-drop" related to the property of RIs?

Third, there have been several proposals claiming that an RI analogue stage could be found even in pro-drop languages. Sano (1995, 1999), for example, has conducted a detailed longitudinal study of three Japanese-speaking children, Toshi (2;3–2;8), Ken (2;8–2;10) and Masanori (2;4), to see if non-finite forms are produced in main clauses. The verb forms he examined are exemplified in (9): the preverbal (*Renyōkei*) form, *-i*, in (9a), the Irrealis (*Mizenkei)* form, *-a*, in (9b), and the Conjunctive form, *-te*, in (9c).

(9) a. *Taroo ga kore ni hair-i-ta-i* (*koto*).
NOM this to enter-(Preverbal)-want-PRS (fact)
'Taro wants to enter into this.'

b. *Taroo ga kore ni hair-a-na-i* (*koto*).
NOM this to enter-(Irrealis)-NEG-PRS (fact)
'Taro does not enter into this.'

c. *Taroo ga kore ni hait-te, Ziroo ga are ni hair-u.*
NOM this to enter-(Conjunctive) NOM that to enter-PRS
'(While) Taro enters into this, Jiro enters into that.'

As shown in Table 2, the Preverbal *-i*, the Irrealis *-a*, and the Conjunctive *-te* were not produced as a main verb by these children, though these forms were produced in non-root contexts, i.e., under finite auxiliary predicates.

**Table 2:** Inflection of main verbs in affirmative declarative root clause (Sano 1999)

| | Non-past-*(r)u* | Past-*ta* | Preverbal | Irrealis | Conjunctive |
|---|---|---|---|---|---|
| Toshi (2;3–2;8) | 288 | 84 | 0 | 0 | 1 (0.2%) |
| Ken (2;8–2;10) | 111 | 175 | 0 | 1 (0.3%) | 0 |
| Masanori (2;4) | 138 | 50 | 0 | 0 | 0 |

Based on data analysis, Sano (1995, 1999) concludes that children at two years of age, who would be in the RI stage in some other languages, do not produce non-finite verbal forms, and hence, there is no RI stage in child Japanese.

Kato et al. (2003) support Sano's conclusion. Pointing out that bare verb stems without tense morphemes are not allowed in adult Japanese, they predict that an RI would have either the present- or the past-tense form. They analyze the corpus of two Japanese-speaking children, Ryo (2;0–3;0) and Tai (2;0–2;9), and find that neither of these forms is overused. Their results are given in Table 3 and Table 4.

**Table 3:** Number of past- or present-tense verbal form in Ryo's corpus (Kato et al. 2003)

| | Past-tense verb forms | Present-tense verb forms |
|---|---|---|
| Correct form | 476 | 761 |
| Erroneous form | 7 | 4 |
| Unclear | 2 | 5 |
| Total | 485 | 770 |

**Table 4:** Number of past- or present-tense verbal form in Tai's corpus (Kato et al. 2003)

| | Past-tense verb forms | Present-tense verb forms |
|---|---|---|
| Correct form | 787 | 1667 |
| Erroneous form | 3 | 15 |
| Unclear | 0 | 14 |
| Total | 790 | 1696 |

As shown above, few erroneous verbal forms are found. Both of the two-year-old children produced present- and past-tense forms in appropriate contexts. Hence, Kato et al. (2003) conclude that an RI stage is not found in child Japanese.

In this chapter, we address two questions: (i) What is RI (analogue) stage? And (ii) what does it mean that verbs are not marked for tense agreement at an early stage of grammar acquisition? We argue that Japanese-speaking children do go through the RI analogue stage, and it is the stage where Tense Phrase is either truncated or Tense/Complementizer elements are jointly, not separately, projected in one node as a T-C head. Non-finite verbs in finite (root) contexts are common in the linguistic production of very young child across languages, but the early verbal forms in child languages reflect the core morphological properties of the adult grammar.[3] We will argue that (i) there is a Very Early Non-Finite Verb Stage in Japanese, (ii) the forms in question are the past-tense form *V-ta* and bare onomatopoeia/mimetics,

3 This analysis does not contradict the descriptive findings reported in Sano (1995) and Kato et al. (2003). Rather, our studies are consistent with their results: Erroneous non-finite verb forms are produced not by two-year-olds, but by much younger children.

(iii) the stage occurs much earlier than Root infinitives in European languages, i.e., even at one year of age, and (iv) the form is initially (around 1;6–1;7) used 100% of the time in the full range of environments.[4]

# 3 Root infinitive analogues in Japanese

## 3.1 Verb forms in adult Japanese and Stem Parameter (Hyams 1986, 2008)

Before we go into RI analogues in child Japanese, let us briefly explain the Japanese verbal conjugation system. In adult Japanese, the bare stems of the verbs cannot appear without tense or aspect morphemes, as shown in (10).

| | | | |
|---|---|---|---|
| (10) | a. | **tabe* | 'to eat' |
| | b. | *tabe*-***ta*** | 'ate' (past/ perfect) |
| | c. | *tabe*-***ru*** | 'eat' (present/ future) |
| | d. | *tabe*-***te (i)ru*** | 'is eating/ have eaten' (present progressive/result state) |
| | e. | *tabe*-***te (i)ta*** | 'was eating/ had eaten' (past progressive/perfect) |
| | f. | *tabe*-***tyatta*** | 'have eaten' (perfective) |
| | g. | *tabe*-***te*** | 'please eat' (request) |

As in (10), the verb stem, *tabe* 'to eat,' itself is not allowed in Japanese. Some morpheme must attach to the verb stem as shown in (10a). The stem is followed by the past tense morpheme -*ta* in (10b), and the present tense morpheme -*ru* in (10c). In (10d), the aspect morpheme -*te i*-, which has either progressive or perfect interpretation, is attached to the verb stem, and it is followed by the present tense morpheme -*ru* to refer to a present progressive event or a result state. In (10e), the past tense morpheme -*ta* attaches to the aspect form, and the form has an either a past progressive or a past perfect interpretation. In (10f), the verb stem is followed by the perfective morpheme -*tyatta*, and in (10g), by the request morpheme -*te*.

**4** See Murasugi and Fuji (2008) for the supporting evidence for Phillips' (1995) insight that the merge of the verb and inflection is not available at the RI Stage. See also Sawada, Murasugi and Fuji (2009) and Sawada and Murasugi (2011) for the report that Japanese-speaking children produce so-called 'the erroneous genitive subjects' (like *Emi-tyan-no* (Emi's) in *Emi-tyan no yattikiru* (Emi will do it) ) at around the age of two just like English-speaking children do (like *my* in *My want one*). We conjecture that this stage is the stage of Optional Infinitives (or a typical RI stage in European languages) where TP is projected, but the features in Tense are underspecified (rather than fully specified). In other words, such forms as V–*ta* form and bare onomatopoeia/mimetics are used as RI analogues when the Tense Phrase is either truncated or Tense/Complementizer elements are jointly, not separately, projected in one node as a T-C head (at around the age of one); while erroneous genitive (and dative) subjects (Murasugi and Watanabe 2009) are optionally used when the features in T are underspecified (at around the age of two) in Japanese.

In fact, whether or not the verb stem can stand by itself without bound morphemes seems to show variation across languages. The *pro*-drop languages, such as Italian or Japanese, seem to share the property that the stem cannot stand by itself. According to Hyams (1986, 2008), languages are parameterized (the Stem Parameter) with respect to whether or not their verbal stem constitutes a well-formed word. For example, as shown in Table 5, in English, a verbal stem, *speak*, is a well-formed word and can stand on its own as a stem. However, in Italian, as shown in Table 6, a verbal stem, *parl-* 'to speak,' is ill-formed. Without any agreement morphemes, the stem of the verb cannot appear in Italian.

**Table 5:** Italian *parl-* (to speak)

| | Singular | Plural |
|---|---|---|
| 1p | *-o* | *-iamo* |
| 2p | *-i* | *-ate* |
| 3p | *-a* | *-anno* |

**Table 6:** English *speak*

| | Singular | Plural |
|---|---|---|
| 1p | – | – |
| 2p | – | – |
| 3p | *-s* | – |

(Hyams 1986)

According to Hyams (1986, 2008), inflectional morphology in a language like Italian represents a "core" property of the language, and it is closely related to the setting of a particular parameter. On the other hand, in English, the Stem Parameter specifies that verbs are uninflected and so the acquisition of the 3rd person, past tense, and progressive morphemes represents a departure from the core grammar of English. This proposal is confirmed by the fact that English-speaking children acquire those morphemes late (Brown 1973, among others), whereas Italian-speaking children acquire verbal inflection relatively very early (Hyams 1986).

Assuming the Stem Parameter, Murasugi, Fuji and Hashimoto (2007) propose that children acquiring [-bare stem] languages produce RI analogues, since the bare stem itself is not a well-formed word in those languages. Japanese-speaking children attach a past tense morpheme *ta* to the verb stem for volition and irrealis meaning as well as for past/perfect events, and the typical properties of RIs listed in (6) are also observed with the verb + *ta* form, and hence, the *V-ta* form is a RI analogue in Japanese.

## 3.2 V-ta forms as root infinitive analogues (RIAs)[5]

In this section, based on the analysis of the longitudinal and observational data of Yuta and corpus analysis of the longitudinal data from Sumihare (Noji 1973–1977, also available in the CHILDES), we show that Japanese-speaking children choose the past tense *V-ta* form as RI analogues, which show some parallel properties with RIs. Importantly, *V-ta* form is initially used 100% of the time with various meanings.

Sumihare and Yuta used *V-ta* form for volition and request as in (11) and (12). This indicates that the RI analogues in Japanese have the Modal Reference Effects just like other languages. First, let's observe Sumihare's data in (11).

(11) a. #*Atti i-ta* (1;6) (adult : volition/ request *ik-u/ik-e*)
there go-PST
'(I) go there / (You) go there.'

b. #*Atti. Atti i-ta* (1;6) (adult :volition/request *ik-u/ik-e*)
there there go-PST
'(I) go there / (You) go there.'

c. #*Sii si-ta* (1;7) (adult : volition *si-tai*)
pee do-PST
'(I) want to pee.'

d. #*Sii si-ta-naa* (1;7) (adult : volition *si-tai*)
pee do-PST-Mood
'(I) want to pee.'

e. #*Baba pai-ta* (1;8) (adult: request *si-te*)
muddy discard-PST
'Please throw (it) away.'
(Murasugi, Fuji and Hashimoto 2007; Murasugi and Fuji 2008, 2009)

In (11a), Sumihare intended to mean 'I want to go there,' or '<u>You</u> go there.' According to Sumihare's father (Noji 1973–1977), he went out with Sumihare, with Sumihare on his back. The father tried to go back home, but Sumihare pointed to a different direction and angrily uttered *atti i-<u>ta</u>* 'there go-PST.' (11b) is a similar example. It is described that Sumihare produced like this when he wanted to go somewhere. In (11c) and (11d), when he wanted to pee, Sumihare uttered *sii si-<u>ta</u>*, using an onomatopoetic expression, *sii*, which means 'to pee.' In adult grammar, the form should be *si-tai* 'want to do,' but Sumihare used the past-tense *ta*-form. In 0, *ta* is attached to

5 See Murasugi, Fuji, and Hashimoto (2007), Murasugi and Fuji (2008, 2009), Murasugi (2009a, b), Nakatani and Murasugi (2009), Murasugi, Nakatani and Fuji (2009), and Murasugi and Nakatani (to appear) for details.

another onomatopoetic expression, *pai*, which means throw away. The situation was that Sumihare had a potato in his hands, and asked his mother to remove mud from the potato. In this context, the request V + *te* form should be used, but *V-ta* form is used instead.

The exactly parallel phenomenon was found with another Japanese-speaking child, Yuta, as shown in (12) (Nakatani and Murasugi 2009).

(12) a. *#Ai-ta.* *Ai-ta* (1;7) (adult: volition/request *ake-ru/ake-te*)
open-PST open-PST
'(I) want to open (the cabinet) / (You) open (the cabinet).'

b. *#Hai-ta.* *Hai-ta* (1;7) (adult: volition/request *hak-u/hak-ase-te*)
put on PST put on-PST
'(I) want to wear (the shoes) / (You) put (the shoes) on (me)

c. *#Hait-ta.* *Hait-ta* (1;7) (adult: volition/request *ire-ru/ire-te*)
enter-PST enter-PST
'(I) want to put (this notebook in this bag) /
(You) put (this notebook in this bag).'

d. *#Tot-ta* (1;7) (adult: volition/request *to-ru/to-tte*)
take-PST
'(I) want to take (the soap) / (You) take (the soap).'
(Nakatani and Murasugi 2009)

In (12a), Yuta used the past tense *V-ta* form, when he wanted to open the cabinet or he wanted to ask his grandmother to open the cabinet. In this context, he should have used the present form, *ake-ru*, or the imperative form, *ake-te*, but instead, he produced the past tense *ta* form. In (12b), he used *hai-ta*, *V-ta* form, when he wanted to wear shoes or he wanted to ask his grandmother to put shoes on him in order to go out. In (12c), Yuta produced *hait-ta*, intending to mean 'I want to put this notebook into this bag', or 'You put this notebook in this bag.' He used *V-ta* form to express his volition or request. Lastly in (12d), *tot-ta* was produced instead of the present form, *to-ru*, or the imperative form, *tot-te*, intending to mean 'I want to take the soap', or 'You take the soap', since Yuta could not reach the soap that he wanted to play with. The data shown above indicate the typical properties associated with RIs, i.e., Modal Reference Effects stated in (6) that have been found in European RIs.

*V-ta* form is used not only for the intentional meaning (volition and request), but also for the extensional meaning (progressive and result state), as stated in (6). It is used instead of the correct aspectual form, such as V + *teiru* and V + *teita*, which have either progressive or result state interpretations. Some examples taken from the Sumihare Corpus are given in (13).

(13) a. #*Baba* *tui-ta* (1;6) (adult: result *tui-te iru*)
thread stick-PST
'The thread stuck (to my finger).'

b. #*Sii* *si-ta* (1;6) (adult: progressive *sii-si-te iru*)
pee do-PST
'(She) is peeing.'

c. #*Buu* *maimai-ta* (1;10) (adult: progressive *si-te iru*)
plane round-PST
'A plane is going round.'

d. #*Akatyan* *gaaze* *oti-ta* (1;11) (adult: result *oti-te i-ta*)
baby gauze drop-PST
'Baby's gauze was on (the floor).'
(Murasugi, Fuji and Hashimoto 2007; Murasugi and Fuji 2008, 2009)

In (13a), Sumihare found a thread on his finger, and intended to inform his mother of this. In this context, an aspectual morpheme, *teiru*, should be attached to the verb stem, but Sumihare uttered *tui-ta*, using *V-ta* form. In (13b), Sumihare employed *V-ta* form instead of V + *teiru* form for the progressive event where one of his friends was peeing. In (13c), he saw a plane flying around and wanted to explain the situation. He used the onomatopoetic expression *maimai*, which means something goes around, and attached the past tense morpheme *ta* to it, instead of the progressive *teiru*. In (13d), he found a baby's gauze towel on the floor and picked it up. In this context, the past perfect ending *teita* should have been used, yielding the form *oti-teita*, but instead he uttered *oti-ta*.

The longitudinal study also found that Yuta used the *V-ta* form instead of the V + *teiru* form for the progressive and result state when he was a late one-year-old, just as Sumihare had done.

(14) a. #*Tui-ta* (1;3) (adult: result *tui-te iru*)
on-PST
'(The light) is on.'

b. #*Oti-ta* *otyoto* *oti-ta* (1;7) (adult: progressive *otosi-te iru*)
drop-PST outside drop-PST
'(I) am dropping (this doll) outside.'

c. #*Tui-ta* (1;6) (adult: result *tui-te iru*)
stick-PST
'(The rice) stuck (to my hand).'

d. #*Oti-ta* *oti-ta* (1;7) (adult: result *oti-te iru*)
drop-PST drop-PST
'(A case of video tapes) is on (the floor).'
(Nakatani and Murasugi 2009)

As in (14a), *tui-ta* was produced as early as 1;3. *Tui-ta* is one of the very first verbs that he produced, and the verb was employed in *V-ta* form 100% of the time until V + *tyatta* form appeared at 1;6. Yuta uttered *tui-ta* when he was watching the light while lying on the sofa. In (14b), Yuta used *oti-ta* when he dropped a doll outside. He seemed to intend to mean 'I am dropping the doll outside.' In this context, he should have used the aspectual morpheme, *teiru*, but he used *ta* instead. In (14c), he uttered *tui-ta* instead of *tui-teiru* when he found rice on his hand. We analyze this utterance as having the result interpretation because the rice had already been stuck on his hand for a while when he found it. Likewise, in (14d), Yuta also used past tense *ta* form instead of aspectual *teiru* form for result state.

As Murasugi, Fuji and Hashimoto (2007) and Nakatani and Murasugi (2009), among others, point out, T(ense)-related items, such as Nominative Case and copulas, and C-related items are not produced with the non-finite verbs. At this stage, either some of the features in T are underspecified or T projection is truncated, as has been pointed out by many researchers (Rizzi 1993/1994; Wexler 1994, among others).

Then, how about the presence of *wh*-questions at this stage? Interestingly, Crisma's effect is observed in Japanese, even though *wh*-questions in Japanese does not require main verbs to move.[6] As in European languages, Tense- or C-related elements (e.g., complementizers and *wh*-phrases) are not found with the non-finite *-ta* forms, as Figure 1 shows.[7]

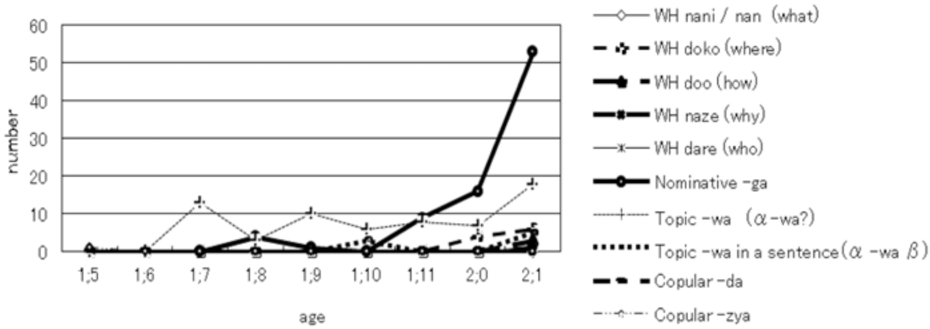



**Figure 1:** Frequency of C-, T- and D-related elements in Sumihare's corpus

These data indicate that the RIs are not merely due to performance deficits of children. Rather, MoodP is active during the Very Early Non-Finite Verb (RIA) Stage, while AspectP and TP are still missing and the head merger inside the verbal projection is still unavailable. Evidence for the lack of Ts (or the jointed T-C heads) found

**6** Nakayama (1997) finds that *wh*-questions start to appear in child production after what we call the Root Infinitive analogue stage.

**7** The topic marker *-wa* was produced at a very early stage, only in the form of NP-*wa*, without ever being followed by verbal predicates.

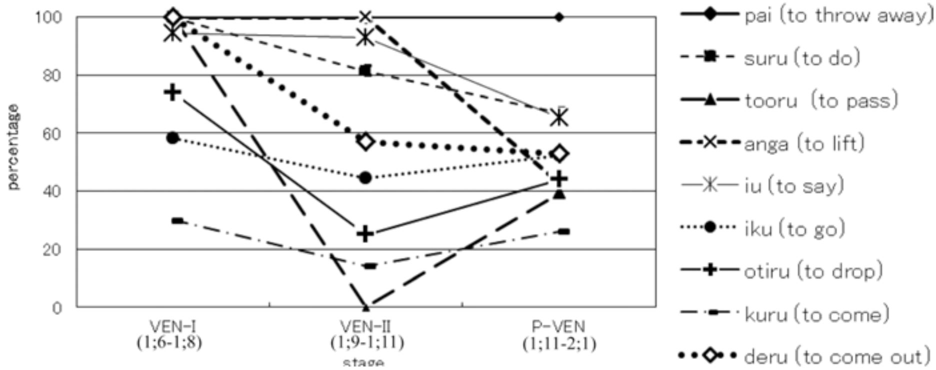



**Figure 2:** Propotion of null topic nominals for each verb in Sumihare's corpus (Murasugi and Fuji 2008, 2009)

in the absence of any other T (or I) elements at the stage in question. Both the Nominative Case marker *-ga* and the finite *da/zya* (the finite *be*, the copula), and the co-occurrence of the tense-related adverbs such as *kinoo* (yesterday) and the RI analogue were not found in Sumihare's corpus, which confirms the possibility that the stage is due to deficits in T (or I) projection (Kishimoto and Murasugi 2013).

Then, what about pro-drop in the subject position? It has been pointed out that RIs tend to co-occur with null subjects more often than finite verbs (i.e., Krämer's effect). As is the case in the acquisition of German (Krämer 1993), Sumihare initially produced null topic nominals (without nominative case marker) frequently with many verbs, though the rate of them was sometimes lower depending on the verb.[8] As shown in Figure 2,[9] the percentage of null topic nominals of speaker-oriented verbs such topic nominals as *pai* (to throw away) or *suru* (to do), where the agent tends to be a speaker (Ego), stays high even after proper inflections (conjugations) appear. On the other hand, subjects (a Topic NP) conveying new information with eventive verbs such as *oti-ru* (to drop) or *ku-ru* (to come) do <u>not</u> tend to be null. This is different from the findings reported in studies of non-null-subject languages, though it should not be surprising given that Japanese is a discourse-*pro* language.[10]

**8** Although verb movement may be involved in the assignment of Nominative Case (Huang 1987; Otani and Whitman 1991), the Nominative Case *-ga* does not appear in the subjects' language at the RI analogue stage. The Nominative Case marker *-ga* first appears around 1;11 for Sumihare.

**9** VEN stands for Very Early Non-Finite Verb Stage, which is divided into two sub-stages: VEN-I is the stage where the *V-ta* form is used almost 100% of the time, and VEN-II is the stage where a modal meaning is realized with the form *tyoodai*. P-VEN stands for Post-Very-Early-Non-Finite Verb Stage.

**10** Kim and Phillips (1998) argue that the overuse of the default mood-inflection "*-e*" in the earliest speech of Korean children parallels the RI in other languages, and report that there is no correlation between the RI analogue form and the number of null subjects produced at this stage. See Murasugi and Fuji (2008) for an argument in favour of a parallelism between the RI analogue stages of Japanese and Korean.

## 3.3 V+ tyatta forms (perfective verb forms) produced by Yuta as surrogate infinitives[11]

Some Japanese-speaking children use another *ta*-form for the RI analogue. Yuta, a Japanese-speaking boy, for example, used V + *tyatta* forms, a perfective form, as RI analogues at the late stage of the very early non-finite verb stage (Nakatani and Murasugi 2009). V + *tyatta* form appeared at 1;6, after the stage when *V-ta* form had been used 100% of the time. Just like *V-ta* form, V + *tyatta* form has the Modal Reference Effect, and shows the properties of RI analogues, as shown in (15).

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (15) | *Kippu* | *kippu* | *kippu* | **ai-ta*. | **ai-ta*. | **ai-tyatta*. | **ai-tyatta*. (1;7) |
| | clips | clips | clips | open-PST | open-PST | open-PERF | open-PERF |

‘(I) want to / (You) open this box of clips.’
(adult: volition/request *ake-tai/ake-te*)

In (15), *ai-ta* and *ai-tyatta* were produced in the same context, and they both had the same intended meaning ‘I want to open this box of clips’, or ‘You open this box’. Hence, V + *tyatta* forms as well as *V-ta* forms are used for volition and request. V + *tyatta* forms are also used for result states.

(16) **Tui-tyatta* (1;7) (adult: result *tui-te iru*)
stick-PERF
‘(The rice) stuck (to my hand).’
(Nakatani and Murasugi 2009)

Yuta uttered (16) when he found rice on his hand. In this context, he should have used the *teiru* form, but instead he used the *tyatta* form. Note that the *V-ta* form, *tui-ta*, was used in a similar context in (14). V + *tyatta* form and *V-ta* form are used in the same manner to express result states.

Interestingly enough, unlike the case of *V-ta* forms, Yuta never used V + *tyatta* form with the meaning of progressive. We analyze that these *tyatta* forms were produced when Yuta found out that *tyatta* is another morpheme that can be attached to the verb stem as well as *ta*, in order to make the stem morphologically well-formed. *Tyatta* is perfective in adult Japanese, but we conjecture that Yuta used these V + *tyatta* forms as non-finite verbs as well as perfective, and this is the first “adult inflection” that the child learned after the stage of non-finite *V-ta* forms used as RI analogues.

11 See Nakatani and Murasugi (2009) and Murasugi and Nakatani (to appear) for details.

## 3.4 Parallels and differences between Sumihare's and Yuta's RI analogue stage

The statistics for the kinds of verbs produced also confirm the predominance of the *V-ta* form forms at the RI analogue stage. The number of instances of each verbal form and the overall proportion of the verbal forms produced by Sumihare between 1;5 and 2;1 are shown in Figures 3 and 4, respectively.

The past tense *V-ta* form is predominantly used until 1;11, and it is used almost 100% of the time at 1;6 and 1;7. The RI analogue stage seems to end at around 1;11, when the present form and other forms appear. Sumihare distinctively used the *tyoodai* 'give me' form between 1;9 and 1;10 in order to express volition and request (e.g., *Pai-tyoodai* 'please throw away'). Interestingly enough, as the frequency of the *tyoodai* form increases, the frequency of *V-ta* forms decreases. This would be

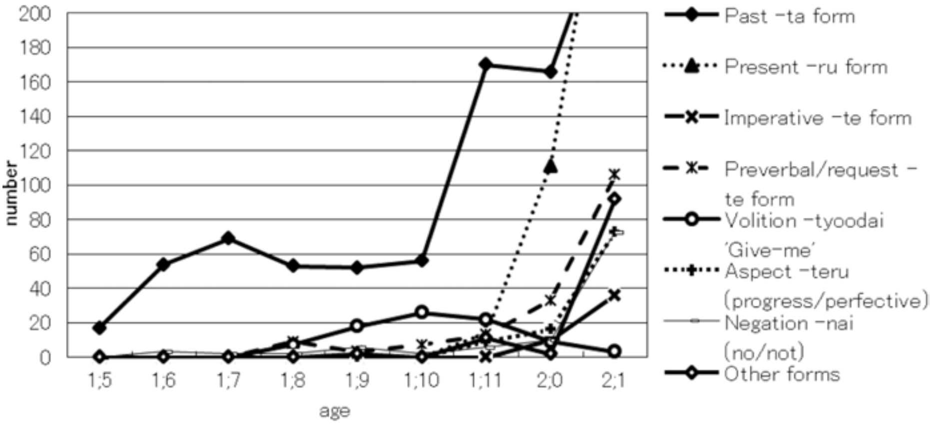



**Figure 3:** Number of instances of verbal forms (Sumihare)

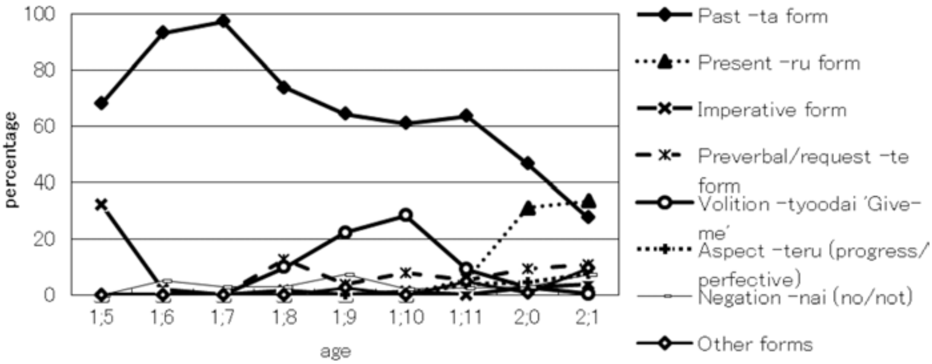



**Figure 4:** Percentage of verbal forms (Sumihare) (Murasugi, Nakatani and Fuji 2009)

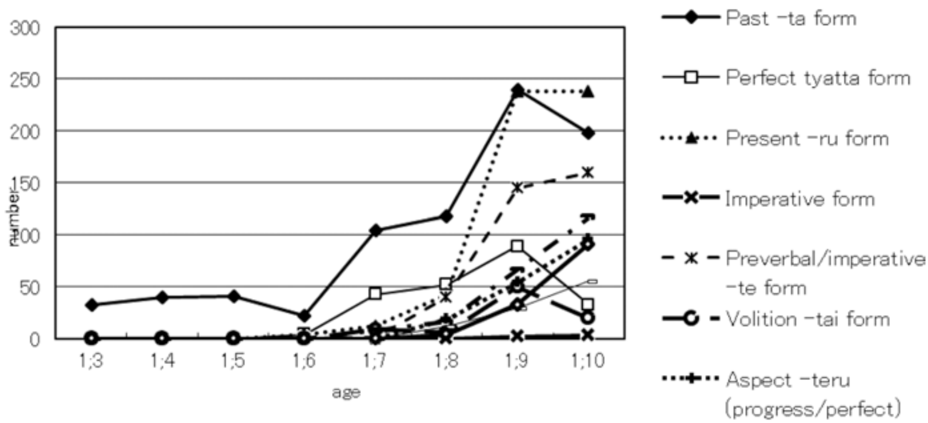



**Figure 5:** Number of instances of verbal forms (Yuta)

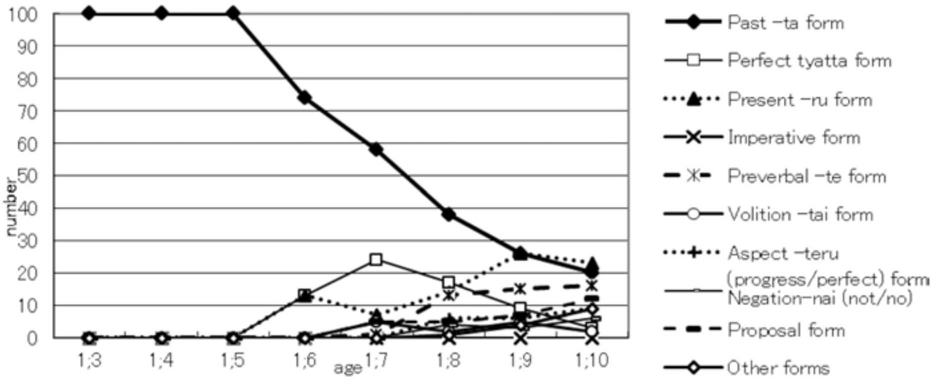



**Figure 6:** Percentage of verbal forms (Yuta) (Murasugi, Nakatani and Fuji 2009)

because volition and request are expressed by *tyoodai* forms, and the *V-ta* form is not used for those meanings anymore.

Importantly, Yuta and Sumihare show parallel curves in the acquisition of verbal conjugations. The results of our analysis of Yuta's production between 1;3 and 1;10 are shown in Figures 5 and 6 (Murasugi, Nakatani and Fuji 2009).

For Yuta, the past tense *V-ta* form appeared at 1;3, and it is predominantly used until 1;8. It is also notable that the perfective V + *tyatta* form appears from 1;6, and that form is the second most predominant until 1;7. Just like Sumihare, the RI analogue stage for Yuta came to an end when the present form and other forms started to appear at around 1;8. On the other hand, Yuta produced the perfective V + *tyatta* form, the volition *V-tai* form, and the propositive form more frequently than Sumihare did.

In this subsection, we argued that child Japanese has a RI analogue stage, and *V-ta* and/or V + *tyatta* is chosen as the RI analogue form. Those forms have the Modal Reference Effects and are predominantly used until other verbal forms appear. In the following subsection, we discuss why Japanese-speaking children go through an RI analogue stage, but not an RI stage.

## 3.5 V-ta as an adult non-finite verb form

Then, why is the *V-ta* form chosen as an RI analogue? Murasugi (2009a) argues that there are several pieces of evidence to indicate that the *V-ta* form is, in fact, the most unmarked non-finite form in adult Japanese.

It is well known that the non-finite *V-ta* form is found in complex NPs in adult Japanese (Teramura 1982; Abe 1993; among others). The past tense morpheme *ta* displays a result state interpretation as well as a past tense interpretation in a relative clause, as in (17).

(17) a. [*boosi-o* *kabut-ta*] *hito*
hat-ACC wear-PST person
(i) 'the person who wore a hat'
(ii) 'the person who is wearing a hat'

b. [*Taroo-ga* *kabut-ta*] *boosi*
NOM wear-PST hat
'the hat which Taro wore'

According to Abe (1993), in (17), the past tense *V-ta* form in a relative clause containing a gap in the subject position denotes not only the past tense reading as in (i), but also the result state reading as in (ii). In (17), the result state reading disappears if a position other than the subject is relativized.

However, Abe (1993) also provides the following examples in (18), which do not contain a subject gap.

(18) a. [*yude-ta*] *tamago*
boil-PST egg
'eggs that are boiled'

b. [*tiisaku* *kit-ta*] *daikon*
small cut-PST radish
'radish cut into small pieces' (Ibid.)

In (18), although the simple past event reading can be detected, the preferred interpretation is the result state.

Furthermore, Murasugi (2009a) shows the non-finite status of *V-ta* form in the non-NP context as well in adult Japanese, discussing such examples as (19) through (21). She argues that the *V-ta* form is used as the strong imperative in Japanese as in (20), just like Italian infinitive (19) in a root clause.

(19) *Partire* *Immediatamente!* (Strong Imperatives in Italian)
go immediately
'Go back (somewhere) immediately!'
(Rizzi 1993/1994)

(20) a. *Kaer-e.*
go back-IMP

b. *Sassa to* *kaet-ta !* *kaet-ta !* (Strong Imperatives in Japanese[12])
immediately go back-PST go back-PST
(Murasugi 2009a)

According to Rizzi (1993/1994), infinitives can appear in a root clause as imperatives in a special context in adult Italian. Similarly, in Japanese, as shown in (20), *V-ta* form, *kaet-ta*, can be used to express the imperative force instead of the imperative form, *kae-re*, as in (20).

In (21), two conjuncts are conjoined by the verbal conjunct *ri* attached to *V-ta* forms, and the form is unspecified for tense.

(21) a. *tabe-ta ri* *non-da ri* *su-ru/-ta.*
eat-PST drink-PST do-PRS/PST
'We eat/ate, and we drink/drank.'

b. *it-ta ri* *ki-ta ri* *de* *taihen* *-da/* *dat-ta.*
go-PST come-PST by troublesome is/ was
'It is/was troublesome (of you) to go back and forth.'
(Murasugi 2009a)

In (22), *V-ta* form is used with irrealis meaning. Murasugi argues that these facts (20b)–(22) indicate that *V-ta* form would be non-finite as well in adult Japanese.

(22) *Mosimo* *watasi ga* *ie o* *tate-ru /-ta* *nara*
if I-NOM house-ACC build-PRS/PST then
*tiisa-na* *ie o* *tate-ru/-ta* (*desyoo*).
small house-ACC build-PRS/PST (would)
'If I built a house, I would build a tiny one.'
(Ibid.)

**12** See also Teramura (1984) and the citation of Kindaichi (1953) there.

Thus, *V-ta* form is the most unmarked surrogate form in both adult and child Japanese, and Japanese-speaking children, even at one year of age, naturally and voluntarily pick up the non-finite form as the default verbal form of their languages, and use it as an RI analogue, as Murasugi (2009a) and Murasugi and Nakatani (to appear) propose.

# 4 The Stem Parameter and the cross-linguistic variation: The surrogate infinitives in [-bare stem] child languages

The discussion so far indicates that very young children speaking Japanese, a typical [-bare stem] language, go through the RI analogue or the Surrogate Infinitive stage. Then how about the other languages sharing the property of [-bare stem]? In this section, based on the descriptions available in previous research, we will argue that children acquiring [-bare stem] languages such as those listed in (23), in fact, undergo an RI analogue stage as well (Murasugi, Nakatani and Fuji 2009).

(23) Child languages that have surrogate forms as root infinitive analogues:
Kuwaiti Arabic, Greek, Romanian, Turkish, Korean, K'iche' Maya, Japanese, among others

The data described in the previous literature can be reinterpreted on independent grounds as showing that children around the age of two who speak [-bare stem] languages attach some morpheme to the verb stem to make a "surrogate form". Since the verb stem itself is not a well-formed word in the language, the very young children pick up the unmarked morpheme in the target language.

Recall here that Dutch has been considered to be a typical RI language, but nevertheless some mysterious phenomena are found. As we saw in (5), repeated below, Wijnen, Kempen and Gillis (2001) report that verbal forms resembling imperatives are found, in addition to the Infinitive forms, at the early two-word stage. If this is the case, then Dutch-speaking children produce the imperative forms as the Surrogate Infinitives as well as the infinitive forms as their first verbs.

(5) "... Starting with the early two-word stage, forms resembling imperatives were discarded from the analyses, as it is unclear whether they are finite or non-finite."
(Wijnen, Kempen and Gillis 2001)

We argued that the fact that more than one type of RI analogue is found in a language is observed not only in Dutch but also in Russian and Italian should

not be labeled coincidental. Very early nonfinite verbs do not necessarily appear in a single form per language.

Then, what about Japanese? Do Japanese-speaking children produce another type of "infinitive" form besides *ta*-forms? We argue that the answer is yes. Very young Japanese-speaking children produce mimetic verbs just at the time when surrogate *ta*-forms are produced at around late one year old. In what follows, we discuss that the mimetic verbs and *ta*-forms are both RI analogues in Japanese.

## 4.1 Onomatopoeic/mimetic verbs in adult Japanese

Japanese is rich in onomatopoeia and mimetic words. They can be used as verbs, nouns, and adverbs in adult Japanese as shown in (24).

(24) a. Mimetic verbs: *giragira suru* (do) 'glare'

b. Onomatopoeic nouns: *wanwan* 'a dog'

c. Onomatopoeic adverbs: *suyasuya nemuru* (sleep) 'sleep peacefully'

Onomatopoeic/mimetic verbs are typically followed by the light verb *suru* 'do'. For example, the mimetic verb *burabura* is followed by the light verb *suru* as in (25). In the structure, *burabura-suru* describes an event 'to walk aimlessly.' Tense and aspect is marked on the light verb as shown in (25).

(25) Onomatopoeic/mimetic verbs followed by the light verb *suru* 'do'

a. *Kooen o* *burabura-su-ru.*
park-ACC MIM-do-PRS
'(I) walk aimlessly in the park.'

b. *Mune ga* *dokidoki-si-tei-ta.*
heart-NOM MIM-do-ASP-PST
'(My) heart was pounding fast.'

Tsujimura (2009) points out that bare onomatopoeia/mimetics, onomatopoeia/mimetics without the light verb *suru*, can be also used as verbs in Japanese. The bare onomatopoeia/mimetic *pisyari* 'shut out' and *shan* 'straighten the back' in (26) are verbs.

(26) Bare oomatopoeia/mimetics (without the light verb *suru*)

a. *Sasaki* *osama ga* *pisyari* *kanpuu* *riree*
king-NOM MIM shutout relay
'The king, Sasaki, shutout a game, and he let his team prevent the opposing team from scoring after several changes of pitchers.'

b. *Sesuzi ga* *syan*
back-NOM MIM
'(He) straightens (his) back.'
(Tsujimura 2009)

In fact, adult bare onomatopoeia/mimetics show the Modal Reference Effects, one of the typical properties of RIs.

(27) a. *Si!* 'Silence!' (Strong Imperative)
b. *Si!* 'Go away!' (Strong Imperative with derogatory connotation)
c. *Sesuzi o syan!* 'Straighten your back!' (Strong Imperative)

The onomatopoeia *si* in (27a) and (27b) can be strong imperatives, meaning 'Silence!' or 'Go away', respectively. *Syan* in (27c) can also be used for an imperative, meaning 'Straighten your back!' This is exactly parallel with the Italian infinitives and the Japanese *ta*-forms that we discussed in (6) and (19).

Thus, RIs may appear in two forms in adult Japanese, and two parametric values of verb morphology may coexist: [+inflected, -stem] verbs such as *V-ta* form, and [-inflected, +stem] verbs such as bare onomatopoeia/mimetics.

## 4.2 Bare onomatopoeia/mimetics as RIAs

In this section, based on the corpus analysis of CHILDES (Sumihare 0;0~6;0, Noji 1973–1977) and the longitudinal study with a Japanese-speaking child, Yuta (0;1~3;5), we argue that the children do produce bare onomatopoeia/mimetics, in addition to Verb+*ta*, as the very early nonfinite verbs.[13]

The children we observed produced the onomatopoeic verbs and *V-ta* form during the same period when T-related elements such as nominative *ga* and C-related elements such as Complementizer and *wh*-phrases were not found.

Bare onomatopoeic/mimetic verbs and *V-ta* forms were predominantly produced until 1;8 when the fully conjugated verb forms are used as shown in Figure 7. These facts naturally lead us to construct a hypothesis that if the bare onomatopoeia/mimetics show the typical properties of RI analogues, then bare onomatopoeia/mimetics produced along with *ta*-forms are analogues as well.

To begin with, there is the question of whether or not children use nominal onomatopoeia/mimetics and verbal onomatopoeia/mimetics distinctively just like adults do. Examples in (28) show that Japanese-speaking children, in fact, used the onomatopoeia/mimetics distinctively.

**13** See Murasugi and Nakatani (to appear) for details.

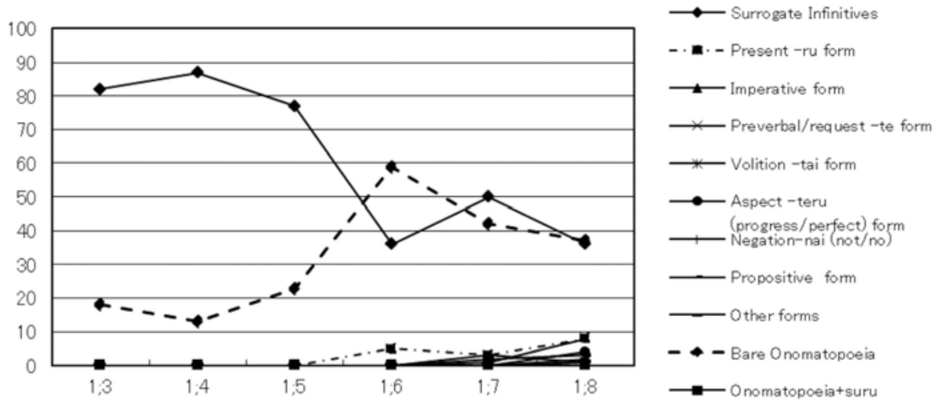



**Figure 7:** Proportion of verbal forms (Yuta) (Murasugi and Nakatani, to appear)

(28) a. *Buu* *it-ta.* *Atti* *it-ta* (S:1;5) [nominal MIM]
MIM go-PST there go-PST
'A three-wheeler went by that way.'

b. *tittyai* *buu buu,* *tittyai* *buu buu* (Y:1;8) [nominal MIM]
small MIM small MIM
'a small car'

c. *dadadadadadadada* (Y:1;6) [looking at shinkansen] [verbal MIM]
MIM
'Shinkansen, a bullet train, is running extremely fast.'

d. *toon-naa* (S:1;7) [S falls down and hits the head.] [verbal MIM]
MIM-mood
'(I) fell down.'

Nominal onomatopoeia/mimetics are exemplified in (28a) and (28b). An onomatopoeia *buu* in (28a), for example, refers to a three-wheeler, which is the subject of the verb *it-ta* 'went'. *Buu buu* in (28b) modified by the adjective *tittyai* 'small' refers to a car. In contrast, in (28c), Yuta produced *dadadadadadadada* when he saw a bullet train, *shinkansen*, which runs very fast. Note here that at this stage, he referred to *shinkansen* as '*shinkantan*' always; he used *dadadadadadadada* only for the on-going action of *shinkansen*. In contrast, onomatopoeia produced by Sumihare were sometimes directly followed by the sentence-ending mood marker *na* to emphasize empathy as in (28d).

The difference between nominal onomatopoeia/mimetics and verbal onomatopoeia/mimetics is also found in the variation of form. The verbal onomatopoeia *buu*, for instance, has variation in its form. Typically, the onomatopoeia used as

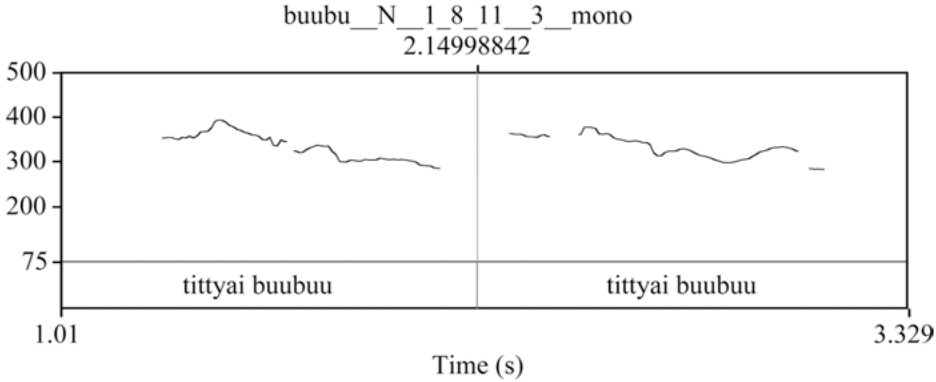



**Figure 8:** Pitch contour of nominal onomatopoeia/mimetic: *buu* (Y:1;8)

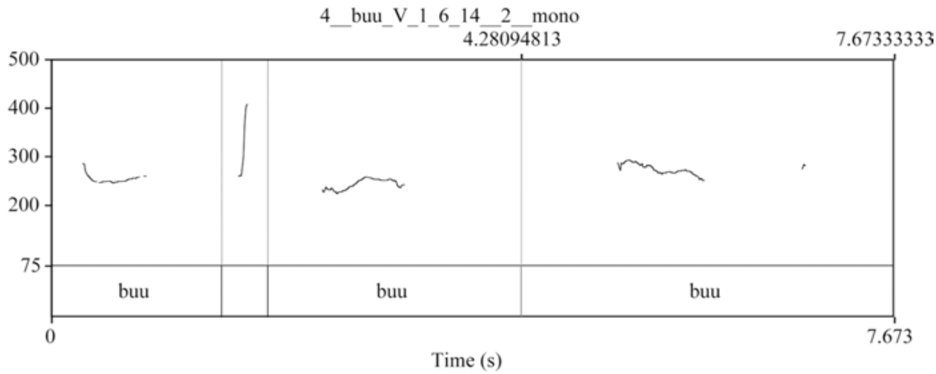



**Figure 9:** Pitch contour of verbal onomatopoeia/mimetic: *buu* (Y:1;6)

verbs are repeatedly pronounced as in *bubuu*, *buu bububuu buu buu*. The observer, Tomomi Nakatani, based on the analysis of the context the onomatopoeia are used, states that the repetition of the onomatopoeia seems to add an adverbial meaning (e.g., fast) to the verbal meaning (e.g., the car runs). Nominal onomatopoeia, such as *wanwan* (a dog), on the other hand, do not have such variation in their form.

Another difference between nominal and verbal onomatopoeia is found in their pitch contours. We used PRAAT (Boersma and Weekink 2005) to measure the pitch contour of each onomatopoeia we collected in the longitudinal study. Figures 8 and 9 show that the nominal *buubuu* and the verbal *buu* are distinct in their pitch accents. A marked fall in pitch is observable in the nominal *buubuu* while the verbal *buu* has flat or rising intonation. Such patterns are also observed in the contrast between the nominal *byuu* and the verbal *byuu*, as shown in Figures 10 and 11.

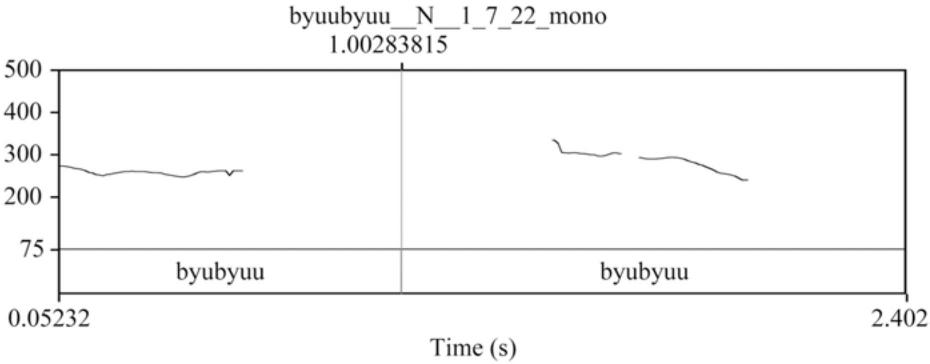



**Figure 10:** Pitch contour of nominal onomatopoeia/mimetic: *byuu* (Y:1;7)

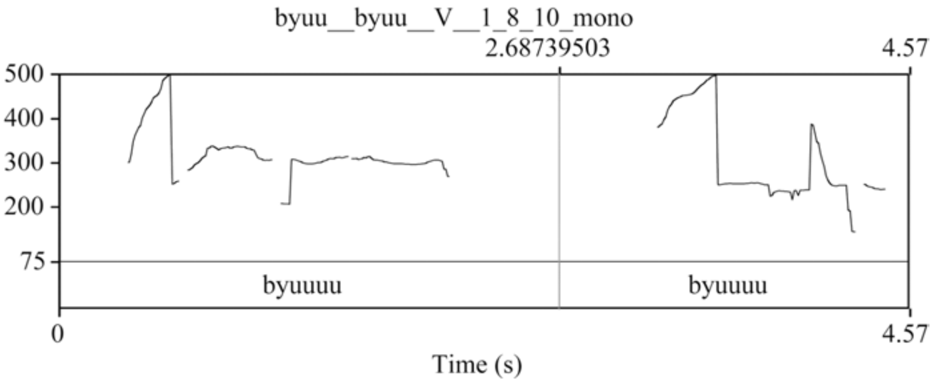



**Figure 11:** Pitch contour of verbal onomatopoeia/mimetic: *byuu* (Y:1;8) (Murasugi and Nakatani, to appear)

Then, are bare onomaopoeia/mimetics associated with the typical properties of RI analogues given in (6)? First, Modal Reference Effects are, in fact, found with Bare onomatopeia/mimetics as shown in (29), just like the typical RI analogues.

(29) Modal Reference Effects of Onomatopoeic RI analogues

a. *baba* *pai-ta* (S:1;8) [S wants mother to remove the dirt on a potato.]
dirt MIM-PST
'(You) remove the dirt.'

b. *odenti* *pai-na* (S:1:10) [trying to take off his gown]
gown MIM-SFP
'(I want to) take off (my) gown.'

c. *buu, buu, buu* (Y:1;6)
MIM
[Y wants grandmother to move the chair that he was sitting on.]
'(You) move (the chair).'

d. *byuuuu, byuuuu* (Y:1;8) [Yuta wants his mother to draw a picture.]
MIM
'(You) draw a picture.'

In (29a), Sumihare produced *baba pai-ta* (mimetic *pai* followed by past tense morpheme *ta*) to ask his mother to remove dirt on a potato. This expression expresses volition and request, but not a past event. (29c) also indicates that *buu buu buu* is produced to order someone to move the chair.

Aspect Effects given in (6) are also found. Bare onomatopoeia were used for progressive and resultative aspect with extensional meaning.

(30) Aspect Effects of onomatopoeic RI analogues

a. *tonton tonton* (S:1;6) [running after children trotting happily]
MIM
'(I) am running.' (progressive)

b. *omoti* *tonton-naa* (S:1;9) [watching rice-cake making]
rice-cake MIM-Sentence-final Particle
'(They are) making rice-cake.' (progressive)

c. *Gasyaan* (Y:1;6) [looking at the broken bowl]
MIM
'This bowl is broken.' (result state)

d. *Pooi* (Y:1;6) [looking at grandfather taking out the trash]
MIM
'Grandpa is taking out the trash.' (progressive)

(30a) shows that Sumihare produced *tonton tonton* to express progressive aspect, and (30c) shows that Yuta produced *gasyaan* to express resultative aspect of a broken bowl, but not to refer to the bowl itself. The bowl itself was never referred to as *gasyaan* by the child in the longitudinal study (Nakatani and Murasugi 2009).

Just like typical RI and RI analogues, bare omatopoeia/mimetics we analyzed as verbs based on analysis with PRAAT are eventive. 100 percent of bare onomatopoeia/mimetics produced by Yuta (1;3~1;8) were eventive (Murasugi and Nakatani in press).

(31) Eventive constraint of onomatopoeic RI analogues

a. Bare onomatopoeia (Sumihare):
*pai* 'remove/take off', *sii* 'pee', *maimai* 'screw',
*toon* 'fall down/drop', *tonton* 'hit/run'

b. Bare onomatopoeia (Yuta):
*buu* 'move', *poi* 'throw a thing', *byuu* 'draw', *jaa* 'pour',
*dada* 'run fast', *biribiri* 'tear', *bibi* 'zip'

In summary, bare onomatopoeia/mimetics in child Japanese shows MREs, Aspect Effects, and the Eventivity constraint. The analysis given above naturally leads us to conclude that there are RI analogues in Japanese. Bare onomatopoeia/ mimetics are, unlike Surrogate Infinitives, followed by no functional elements. Bare onomatopoeia are, rather, like the bare verbs without functional elements that Swahili-speaking children produce as RI analogues as given in (32).

(32) RI analogues in Swahili

| | | | | | | |
|---|---|---|---|---|---|---|
| Child: | *Ø* | *-Ø* | *-ka* | *-a* | *hapa* | (2;3) |
| Adult: | *a* | *-na* | *-ka* | *-a* | *hapa* | |
| | SA3s | -PRS | live | IND | here | |

'She lives here.'
(Deen 2002)

Bare verbs as RI analogues are also observed in other [+bare stem] languages. Inuktitut-speaking children produce bare verbs which are ungrammatical in their target language, as shown in (33).

(33) RI analogues in Inuktitut

a.

| | | | |
|---|---|---|---|
| *Kuapa* | *liar* | *uma* | *paa*. (Elijah 2;0) |
| kuapa | -liaq | -guma | -paaq |
| coop | -VZ.GO.TO | -MODAL.SIFFOX | -OH.HOW.I |

'Oh, how I want to go to the co-op.'

b.

| | | |
|---|---|---|
| *Kuapa* | *lia*. (Elijah 2;0) | |
| kuapa | -liaq | -φ |
| coop | -VZ.GO.TO | -NO.INFL |

'(I want to) go to the co-op.'
(Swift and Allen 2002)

Swift and Allen (2002) observe that MREs are found when inflection drops. The child, who could produce the full form in (33a), produced bare verbs omitting the verbal inflection in (33b) when he expressed his strong volition.

The parallel phenomenon is found in Malagasy as well. In (34), the child omitted a morpheme of past tense and "actor trigger" *ni* which is obligatory in adult Malagasy.

(34) RI analogues in Malagasy

a. *Tomany* *za* (Tsiorisoa 2;7)
cry 1SG NOM STR
'I cried.'

b. *Ni* *tomany* *aho* (adult form)
PST AT cry 1SG NOM STR
(Ntelitheos and Manorohanta 2004)

Nonfinite verbs appear as bare onomatopoeia and the *V-ta* form in Japanese. Onomatopoeia and the *V-ta* form can be nonfinite in adult and child Japanese. Even one-year-old children naturally acquire the two parametric values possible in the target language, i.e., [+bare stem] and [-bare stem], and produce the two types of RI analogues as their first verbs.

# 5 Conclusion

Root Infinitives (RIs) are non-finite (infinitival) verbal forms which very young children use in matrix (root) clauses, where such forms are not possible in adult grammar. Whether or not the target language is *pro*-drop, children go through the very early non-finite verb stage. The children's use of non-finite verbs in root contexts is a universal phenomenon, but there are morphological variations associated with the different verbal systems of the target languages. RIs can be infinitives, bare verbs, participles, or certain (surrogate) "finite" forms. Japanese RI analogues are Verb + *ta* form (or *tyatta* form) and onomatopoeia/mimetics.

What children tell us is that there are several types of possible nonfinite verb forms in human languages, and Stem Parameter, a parameter related to the verbal morphology, plays a role in determining the form of the very early non-finite verbs.

Root Infinitives produced by children suggest that they go through a stage at which they speak a language that is like adult grammar in many respects, but one that is also like other languages, in allowing for the sentences without independent T projection. The "tense-less" phrases (or the phrases with the jointed T-C heads) children produce across languages cannot be explained by an experience-dependent account; adults speaking Japanese and English, for example, never produce those. The phenomena found at the intermediate stages of language development are, just like variation among the world's languages (or the set of internalized I-languages), restricted within the range of Universal Grammar innately endowed to human beings.

## Acknowledgments

The research presented here would not have been possible without discussions with colleagues and students involved in the activities of the Center for Linguistics at Nanzan University. I would like to deeply thank Tomomi Nakatani Murai, Chisato Fuji, Tomoko Hashimoto, and Naoko Sawada, who have been working on the topic discussed here together, and allowed me to introduce the results we have obtained. Special thanks also go to Mamoru Saito, Luigi Rizzi, Ken Wexler, Bonnie Schwartz, Kamil Deen, Diane Lillo-Martin, William Snyder, Jonah Lin, Koji Sugisaki, Kensuke Takita, Hideki Kishimoto, Dylan Tsai, K.A. Jayaseelan, R. Amritavalli, and Yuma Iwatani, for sharing insights. Deepest gratitude is due to JJ Nakayama and the anonymous reviewers for their invaluable comments.

The research reported here was supported in part by National Institution for Japanese Language and Linguistics (NINJAL) collaborative research project "Linguistic variations within the confines of the language faculty: A study in Japanese first language acquisition and parametric syntax" (P.I.: Keiko Murasugi), by JSPS Grant-in-Aid to Nanzan University (#20520397 and #26370615, P.I.: Keiko Murasugi) for the study of the Root Infinitives, functional categories and Tense Parameter, and by Nanzan University Pache Research Grant I-A 2012-2014. Sincere thanks go to the Center for Linguistics at Nanzan University for providing all the support needed for our research activity. Thanks also go to the members of the series of the International Symposium (2006-present) of the Cambridge-Connecticut-Hyderabad-Nanzan-Siena-Tsinghua Consortium in Linguistics, which are the base of the present work.

## References

Abe, Yasuaki. 1993. Dethematized subjects and property ascription in Japanese. *Language, information and computation, Proceedings of Asian conference*. 132–144. Seoul: Thaehakea.

Bar-Shalom, Eva and William Snyder. 2001. Descriptive imperatives in child Russian and early correct use of verbal morphology. In Anna H.-J. Do, Laura Domínguez, and Aimee Johansen (eds.), *Proceedings of the 25th annual Boston University conference on language development*, 94–101. Somerville, MA: Cascadilla Press.

Bel, Aurora. 2001. Sujetos nulos y sujetos explícitos en las gramáticas iniciales del castellano y el catalán [Null subjects and explicit subjects in the initial grammars of Castilian and Catalan]. *Revista Española de Lingüística* 31(2). 537–562.

Blom, Elma and Frank Wijnen. 2000. How Dutch children's root infinitives become modal. In S. Catherine Howell, Sarah A. Fish, and Thea Keith-Lucas (eds.), *Proceedings of the 24th annual Boston University conference on language development*, 128–139. Somerville, MA: Cascadilla Press.

Brown, Roger. 1973. *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Chien, Madelaine. 2008. *Is there a root infinitive or root infinitive analogue stage in early Mandarin?* Qualifying Exam Paper, National Tsing Hua University, Taiwan.

Deen, Kamil Ud. 2002. *The acquisition of Nairobi Swahili: The morphosyntax of inflectional prefixes and subjects*. Los Angeles, CA: UCLA dissertation.

Grinstead, John. 1998. *Subjects, sentential negation and imperatives in child Spanish and Catalan*. Los Angeles: UCLA dissertation.

Grinstead, John. 2000. Case, inflection and subject licensing in child Catalan and Spanish. *Journal of Child Language* 27. 119–156.

Guasti, Maria Teresa. 2002. *Language acquisition: The growth of grammar*. Cambridge, MA: MIT Press.

Guasti, Maria Teresa. 2004. *Language acquisition: Growth of grammar*. Cambridge, MA: MIT Press.

Haegeman, Liliane. 1995. Root infinitives, tense, and truncated structures in Dutch. *Language Acquisition* 4. 205–255.

Hoekstra, Tuen and Nina Hyams. 1998. Aspects of root infinitives. *Lingua* 106. 91–112.

Huang, C.-T. James. 1987. Remarks on empty categories in Chinese. *Linguistic Inquiry* 18. 321–337.

Hyams, Nina. 1986. *Language acquisition and the theory of parameters*. Dordrecht: D Reidel Publishing Co.

Hyams, Nina. 2002. Clausal structure in child Greek: A reply to Varlokosta, Vainikka and Rohrbacher and a reanalysis. *The Linguistic Review* 19. 225–270.

Hyams, Nina. 2005. Child non-finite clauses and the mood-aspect connection: Evidence from child Greek. In Paula Kempchinsky and Roumyana Slabakova (eds.), *The syntax, semantics and acquisition of aspect*. 293–316. Dordrecht: Kluwer.

Hyams, Nina. 2008. The acquisition of inflection: A parameter-setting approach. *Language Acquisition* 15. 192–209.

Kato, Sachiko, Yumi Sato, Yukiko Takeda, Ritsuko Miyoshi, Yumi Sakai and Masatoshi Koizumi. 2003. Root infinitives: Nihongo karano kenshō [Root infinitives: From the perspectives of Japanese]. *Tohoku University Linguistics Journal* 12. 113–127.

Kim, Meesook and Colin Phillips. 1998. Complex verb construction in child Korean: Overt markers of covert functional structure. In Annabel Greenhill, Mary Hughes, Heather Littlefield, and Hugh Walsh (eds.), *Proceedings of the 22nd annual Boston University conference on language development*, 430–441. Somerville, MA: Cascadilla Press.

Kindaichi, Haruhiko. 1953. Fuhenka jodōshi no honshitsu (1, 2): Shukanteki hyōgen to kyakkanteki hyōgen no betsu ni tsuite [The essence of uninflecting auxiliaries (1, 2): About the distinction between subjective and objective expressions]. *Kokugo Kokubun* 22(2). 1–18; 22(3). 15–35.

Kishimoto, Hideki and Keiko Murasugi. 2013. Projection of tense and adverbs. Paper presented at the Lexicon Festa, NINJAL, Tokyo, 1 February.

Krämer, Irene. 1993. The licensing of subjects in early child language. In Colin Phillips (ed.), *Papers on case and agreement I*, *MIT working papers in linguistics* 19. 197–212.

Lillo-Martin, Diane and Ronice Müller de Quadros. 2009. Two in one: Evidence for imperatives as the analogue to RIs from ASL and LSB. In Jane Chandlee, Michelle Franchini, Sandy Lord, and Gudrun-Marion Rheiner (eds.), *Proceedings of the 33rd annual Boston University conference on language development*, 302–312. Somerville, MA: Cascadilla Press.

Murasugi, Keiko. 2009a. What Japanese-speaking children's errors tell us about syntax. Paper presented at the Asian Glow VII, English and Foreign Languages University, Hyderabad, India, 28 February.

Murasugi, Keiko. 2009b. The onset of complex NPs in child production. Paper presented at WAFL 6, Nagoya University, 5 September.

Murasugi, Keiko and Chisato Fuji. 2008. Root infinitives: The parallel route that Japanese- and Korean-speaking children step in. Paper presented at Japanese-Korean Linguistics Conference 18, City University of New York, 13 November. (Paper appeared in Japanese and Koran Linguistics 18, CSLI Publications, pp. 3–15, 2011)

Murasugi, Keiko and Chisato Fuji. 2009. Root infinitives in Japanese and the late acquisition of head-movement. *BUCLD* 33 *Proceedings supplement*.

Murasugi, Keiko, Chisato Fuji and Tomoko Hashimoto. 2007. What's acquired later in an agglutinative language. Paper presented at the Asian GLOW VI, Chinese University of Hong Kong, 27 December.

Murasugi, Keiko, Chisato Fuji and Tomoko Hashimoto. 2010. What's acquired later in an agglutinative language. *Nanzan Linguistics* 6. 47–78.

Murasugi, Keiko, Tomomi Nakatani and Chisato Fuji. 2009. The roots of root infinitive analogues: The surrogate verb forms common in adult and child grammars. *BUCLD* 34 Supplement. http://www.ic.nanzan-u.ac.jp/LINGUISTICS/staff/murasugi_keiko/pdf/2010-murasugi_nakatani_fuji.pdf

Murasugi, Keiko and Tomomi Nakatani (to appear) "Three types of 'root infinitives': Theoretical implications from child Japanese" Proceedings of. 20th Japanese/Korean Linguistics. Conference (JK 20) at University of. Oxford.

Murasugi, Keiko and Eriko Watanabe. 2009. Case errors in child Japanese and the implication for the syntactic theory. *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America*. 153–164.

Nakatani, Tomomi and Keiko Murasugi. 2009. Gengo kakutoku ni okeru shusetsu futēshi genshō: Jūdanteki kansatsuteki kenkyū [Root infinitive analogues as non-finite surrogate forms: A longitudinal study of a Japanese-speaking child]. *Academia* 86. 59–94.

Nakayama, Mineharu. 1997. Null subjects and *wh*-questions in child Japanese. Paper presented at the Workshop on First Language Acquisition of East Asian Languages, Cornell University, July.

Noji, Junya. 1973–1977. *Yōjino gengoseikatsuno jittai [The language use in child age]* I–IV. Tokyo: Bunka Hyōron Shuppan.

Ntelitheos, Dimitrios and Cecile Manorohanta. 2006. Default pronouns and root infinitives in Malagasy acquisition. In K. U. Deen, J. Nomura, B. Schulz, and B. D. Schwartz. *The Inaugural Conference on Generative Approaches to Language Acquisition in North America (GALANA)*. Univeristy of Connecticut Occasional Papers in Linguistics #4. MITWPL. 249–260.

Otani, Kazuyo and John Whitman. 1991. V-raising and VP-ellipsis. *Linguistic Inquiry* 22. 345–358.

Radford, Andrew. 1999. Genitive Subjects in Child English. University of Essex. http://privatewww.essex.ac.uk/~radford/PapersPublications/gensubjects.htm

Phillips, Colin. 1995. Syntax at age two: Cross-linguistic differences. *MITWPL* 26. 325–382.

Phillips, Colin. 1996. Root infinitives are finite. In Andy Stringfellow, Dalia Cahana-Amitay, Elizabeth Hughes and Andrea Zukowski (eds.), *Proceedings of the 20th Annual Boston University Conference on Language Development*, 588–599. Somerville, MA: Cascadilla Press.

Rasetti, Lucienne. 2003. *Optional categories in early French syntax: A developmental study of root infinitives and null arguments*. Genève: University of Genève dissertation.

Rizzi, Luigi. 1993/1994. Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition* 3. 371–393.

Salustri, Manola and Nina Hyams. 2003. Is there an analogue to the RI stage in the null subject languages? In Barbara Beachley, Amanda Brown, and Frances Conlin (eds.), *Proceedings of the 27th annual Boston University conference on language development*, 692–703. Somerville, MA: Cascadilla Press.

Salustri, Manola and Nina Hyams. 2006. Looking for the universal core of the RI stage. In Vicenç Torrens and Linda Escobar (eds.), *The acquisition of syntax in Romance languages*. Amsterdam: John Benjamins. 159–182.

Sano, Tetsuya. 1995. *Roots in language acquisition: A comparative study of Japanese and European languages*. Los Angeles, CA: UCLA dissertation.

Sano, Tetsuya. 1999. Verbal inflection in the acquisition of Japanese. http://core-sun.kuis.ac.jp/public/paper/outside/sano2.pdf.

Sano, Tetsuya and Nina Hyams. 1994. Agreement, finiteness, and the development of null arguments. *NELS* 24. 544–558.

Sawada, Naoko and Keiko Murasugi. 2011. A cross-linguistic approach to a genitive subjects: Under-specification of tense in child grammar revisited. Selected Proceedings of the 4th Conference on Generative Approaches to Language Acquisition North America (GALANA 2010). 209–226.

Sawada, Naoko, Murasugi, Keiko, and Fuji, Chisato. 2009. A theoretical approach to the genitive subjects in child Japanese and the specification of tense. *BUCLD* 34 Supplement. http://www.bu.edu/linguistics/BUCLD/Supp34/Sawada_Naoko.pdf.

Swift, Mary D. and Shanley E. M. Allen. 2002. Contexts of verbal inflection dropping in Inuktitut child speech. In Barbora Skarabela, Sarah Fish, and Anna H.-J. Do (eds.), *Proceedings of the 26th annual Boston University conference on language development*, 670–681. Somerville, MA: Cascadilla Press.

Teramura, Hideo. 1982. *Nihongo no shintakkusu to imi [Japanese syntax and meanings]*. Tokyo: Kurosio Publishers.

Teramura, Hideo. 1984. *Nihongo no shintakkusu to imi [Japanese syntax and meanings]* 2. Tokyo: Kurosio Publishers.

Tsujimura, Natsuko. 2009. Onomatope dōshi no imi: Kō kōzō no ichi kōsatsu [Examination of the meaning and argument structure of mimetic verbs], *KLS* 29; *Proceedings of the Thirty-Third Annual Meeting*. 334–343.

Varlokosta, Spyridoula, Anne Vainikka and Bernhard Rohrbacher. 1996. Root infinitives without infinitives. In Andy Stringfellow, Dalia Cahana-Amitay, Elizabeth Hughes and Andrea Zukowski (eds.), *Proceedings of the 20th annual Boston University conference on language development*, 816–827. Somerville, MA: Cascadilla Press.

Wijnen, Frank, Masja Kempen and Steven Gillis. 2001. Root infinitives in Dutch early child language: An effect of input? *Journal of Child Language* 28. 629–660.

Wexler, Kenneth. 1994. Optional infinitives, head movement, and economy of derivation. In David Lightfoot and Norbert Hornstein (eds.), *Verb movement*, 305–350. Cambridge: Cambridge University Press.

Takuya Goro

# 5 Acquisition of scope

## 1 Introduction

In natural languages, scope relationships between logical words do not always correspond to surface linear/hierarchical relationships. Consider the English sentences in (1) and (2), which allow "inverse" scope interpretations where a hierarchically lower quantificational element takes wider scope than a higher one:

(1) *John didn't find someone.*
can mean: 'there is someone that John didn't find'

(2) *Everyone didn't read this book.*
can mean: 'not everyone read this book'

In (1), *someone* in the object position takes scope over negation. In contrast, *everyone* in the subject position of (2) can be interpreted under the scope of negation. The existence of inverse scope interpretations demonstrates that the mapping between surface syntax and semantics is not always simple, and suggests that the mapping system of a natural language encompasses some mechanism that connects the mismatching surface syntactic and semantic representations.

Scope flexibility in natural languages is not an unrestricted phenomenon. Rather, many sentences/constructions in natural languages are scopally unambiguous, allowing only one of the logically possible scope interpretations. Particularly relevant to our concern in this chapter is the existence of various language-specific constraints on scope interpretation. For example, the following construction in Japanese does not show the scope ambiguity of its English counterpart in (4):

(3) *Dareka ga* *dono-sensei mo* *hihan-sita.*
someone NOM every teacher criticize-did
Literally: 'Someone criticized every teacher.'
$\exists >> \forall$ / $*\forall >> \exists$

(4) *Someone criticized every teacher.*
$\exists >> \forall$ / ${}^{OK}\forall >> \exists$

The English sentence in (4) allows an inverse scope interpretation in which the universal quantifier *every* takes scope over the existential quantifier *some*: For every teacher, there is some individual who has criticized him. In Japanese transitive sentences with multiple QP arguments such as (3), the inverse scope interpretation

is disallowed; the object QP cannot take a scope wider than that of the subject QP (e.g., Hoji 1985; Kuno 1973). Thus, the sentence can only mean that there is a specific individual who has criticized every teacher. Based on Huang (1982), this is called *the rigid scope constraint*. Similar scope rigidity has been observed in other East-Asian languages such as Korean (e.g., Beck and Kim 1997; Kim 1989) and Chinese (e.g., Aoun and Li 1989; Huang 1982).

The properties of natural language scope phenomena pose a problem for first language learners, and therefore, for a theory of language acquisition. Since languages and constructions vary with respect to possible scope interpretations, some form of learning must be involved in the mastery of the relevant linguistic knowledge. For example, the interpretive contrast between (3) and (4) suggests that while children acquiring English somehow learn that sentences such as (4) allow inverse scope interpretations, children exposed to Japanese will eventually learn that inverse scope is impossible with sentences such as (3). Given this, a theory of language acquisition must provide an explanation for how first language learners solve this learning problem. In other words, it must determine how children learn both what is possible and impossible in their target language with respect to scope.

This chapter aims to provide an overview of the problems that are involved in the acquisition of scope. First, I will define the task for an acquisition theory related to language-specific constraints on scope interpretation in terms of the learnability approach (e.g., Pinker 1979, 1989; Wexler and Culicover 1980; Baker and McCarthy 1981), and specify what must be uncovered by empirical investigation. I will then review the empirical data found in recent experimental studies, and discuss the consequences of these findings.

## 2 Learnability

To construct an acquisition theory for a given piece/domain of linguistic knowledge, the following components of the acquisition process must be specified: (i) the learner's contribution; (ii) the learner's experience; and (iii) what is learned. Pinker (1989) indicated that when each of those components has a certain characteristic, the acquisition problem becomes a paradox which cannot be explained in a logically reasonable way. I will briefly review each of Pinker's points, and then explain the relationship of each point to the problem of scope acquisition.

### 2.1 Productivity of the learner

Every natural language allows unboundedly many expressions, and the acquisition of language is carried out on the basis of finite numbers of input sentences that children hear from their parents, so language acquisition cannot be a strictly conservative process. Learners cannot simply stick to expressions that they have heard

and must make generalizations that go beyond their finite linguistic experience in order to productively generate new expressions that were not included in the input data. However, the task of language acquisition becomes complicated when a learner's generalization is not appropriately restricted, generating a certain expression X that is not possible in the target language. In such a case, the learner must learn to modify her hypothesized generalization so that it correctly blocks expression X. Given this consideration, our first empirical task is to determine whether children overgenerate scope interpretations. If children are found to be so productive that they generate scope interpretations that are not possible in the target language, then we must ask how they learn to correct their overly permissive generalization. In Section 3, I will review empirical data from experimental studies that bear on this point. Recent studies have found that across languages and constructions children do not appear to be sensitive to language-specific constraints on scope, thus generating scope interpretations that adults do not allow. These data lead us to consider the mechanisms children use to purge their non-adult scope interpretations, and the contributions of their experience to the process.

## 2.2 Obtaining negative evidence from experience

One possibility is that input experience provides some kind of *negative evidence* to the learner, leading her to recognize that expression X is impossible in the target language. One type of negative evidence that has been extensively discussed in the literature is direct negative evidence, i.e., some sort of parental feedback (e.g., correction, disapproval, etc.) to children's utterances. However, the available evidence suggests that direct negative evidence is not systematically provided to children. For example, since the study of Brown and Hanlon (1970), research on child–parent interactions has repeatedly found that the form of parental feedback to children's speech is not contingent on the well-formedness of children's utterances. Given this, Pinker concludes that learners cannot count on direct negative evidence to determine what is impossible in the target language.

With respect to scope interpretation, it seems straightforward to assume that children do not receive direct negative evidence. Direct negative evidence against a particular scope interpretation could only arise when: (i) the child uses a doubly-quantified sentence intending that scope interpretation; (ii) the caretaker correctly identifies the child's intended scope interpretation; and (iii) the caretaker corrects the child in a way that it is clear that the problem is her scope assignment (rather than, for example, the choice of the particular lexical items). Given that parental feedback is highly inconsistent even in the cases where children's errors are much more obvious (i.e., errors in forms, rather than in interpretations), it is extremely unlikely that children encounter such a situation.

A potential surrogate for direct negative evidence in the form of parental feedback is *indirect negative evidence*. Roughly speaking, indirect negative evidence is the absence of input evidence for a certain structure/interpretation. If the learner is able to detect a systematic absence of a particular scope interpretation, then she may be able to infer from her experience that the scope interpretation is not permitted in the target language. In Section 4, I will discuss this possibility. I will argue that the nature of the input data concerning scope interpretations makes it highly unlikely that children rely on indirect negative evidence in learning language-specific constraints on scope. Thus, my general conclusion will be that negative evidence (direct or indirect) does not play a significant role in the acquisition of scope. If this is the case, we need an alternative way to explain how children correct their non-adult hypotheses regarding scope.

## 2.3 Arbitrariness of the constraint

It is possible that the impossibility of expression X is a consequence of some general property of the language. In such a case, the learner does not have to be directly told that X is impossible; by learning the general property, she should know to expunge X from her language. However, it is also possible that the constraint blocking the generation of X is *arbitrary* in the sense that the impossibility of X is not related to any other property of the grammar. For example, the impossibility of a particular scope interpretation with a certain quantificational element may turn out to be a purely idiosyncratic feature of that lexical item. In such a case, the learner cannot avoid the problem of "unlearning" by learning other aspects of the target grammar. Once the learner makes an overly permissive generalization that allows the scope interpretation, then she must be able to find evidence against the interpretation in the input data.

Pinker points out that a learnability paradox arises when an acquisition task has the following three characteristics: (i) *productivity* – the learner productively generates new expressions in such a way that some ungrammatical expressions are also generated; (ii) *no negative evidence* – experience does not provide the learner with any form of negative evidence against the ungrammatical expressions; and (iii) *arbitrariness* – the ungrammaticality of those expressions is not predictable from other generalizations that can be made on the basis of observable properties of the language, and therefore, the learner needs direct evidence showing that the expressions are not possible in the language. With all these aspects, the acquisition of that piece of linguistic knowledge resists a logical learnability account. The learner makes a mistake that must be corrected by the time she becomes an adult, but there is no reasonable way to explain how the correction occurs. Accordingly, a theory of language acquisition must deny at least one of the three components.

The same kind of learnability paradox can arise in the acquisition of language-specific constraints on scope interpretation, such as (i) children's grammar generates non-adult scope interpretations in addition to scope interpretations that are possible in the target language; (ii) input data do not provide any kind of negative evidence against children's non-adult scope interpretations; and (iii) the relevant constraint is arbitrary in the sense that the impossibility of those scope interpretations is not related to any other properties of the language. Therefore, a theory of scope acquisition must deny at least one of the three components. In Section 5, I will review theoretical accounts that attempt to deny the third component: arbitrariness. These accounts are motivated by the empirical findings reviewed in Section 3, which suggest that children tend to over-generate scope interpretations, ignoring certain language-specific constraints on scope interpretation. These empirical data, combined with consideration of the unreliable nature of negative evidence, force us to come up with an account that derives the effect of the relevant scope constraint from other properties of the grammar. Section 6, by contrast, provides an overview of a case where experimental data deny the first component, productivity. Within a restricted set of quantificational elements, children do not over-generate scope interpretations. Rather, they restrict their scope interpretations in a way that any mismatch with adult grammar can be detected on the basis of positive evidence alone. Thus, first language learners have multiple ways to solve the problem of scope acquisition, and utilize different learning strategies for different scope phenomena. In some cases, children are initially over-productive with scope interpretations, and later learn to expunge some of their interpretations through further learning. In other cases, they narrowly restrict their scope interpretations, and conservatively modify their grammar according to input evidence.

# 3 Productivity and conservatism

## 3.1 Conservative learning approach

A learnability paradox arises when the learner makes an over-generating generalization that cannot be falsified by input evidence. One way to avoid being stuck with over-generating generalizations is to avoid making such generalizations; if you do not make a mistake in the first place, you do not have to correct your mistake. In literature, this idea has usually been implemented in some form of *conservative learning* algorithm. Such a conservative algorithm forces the learner to choose the most restrictive generalization and to hold that generalization until positive evidence shows that it is too restrictive.

This kind of learning mechanism has widely been assumed in various approaches to language acquisition. Within the Principles and Parameters approach (e.g., Chomsky 1981, 1986), the idea is often implemented in *the Subset Principle* for

parameter setting. The Subset Principle forces the learner to choose the parameter value that yields the most restrictive grammar, i.e., the grammar that generates the smallest subset of sentences, until positive evidence proves that the parameter setting cannot generate possible sentences in the language (e.g., Berwick 1985; Clark 1992; Fodor 1992, 1994; Manzini and Wexler 1987; Roeper and de Villiers 1987; Wexler 1993). In addition to the subset principle for syntactic acquisition, some studies on semantics acquisition have proposed the Semantic Subset principle, which states that children assume – as a default – the scope interpretation that yields the narrowest truth conditions (i.e., the interpretation that makes the sentence true in the fewest possible situations; Crain, Ni, and Conway 1994; Goro and Akiba 2004; Goro 2007; Jing et al. 2005). I will return to the Semantic Subset Principle in Section 6.

In the beginning of the 2000s, experimental findings led some researchers to argue that young children are indeed conservative learners of possible scope interpretations. The relevant findings were reported in Musolino, Crain, and Thornton (2000), which claimed that young children's scope interpretations are restricted to those that match surface word orders.[1] In their Truth Value Judgment Task experiments, Musolino, Crain, and Thornton found that young children often failed to assign inverse scope readings to test sentences such as those in (5) and (6):

(5) *The detective didn't find someone/some guy.*

(6) *Every horse didn't jump over the fence.*

In one set of experiments, English-speaking children were presented with sentence (5) as a description of a situation where the detective found two of his friends but missed the third one. Under the surface scope interpretation of (5) (i.e., NOT >> ∃), the sentence is false, because the detective did find someone. The inverse scope interpretation (i.e., ∃ >> NOT), however, makes the test sentence true, because there is indeed someone that the detective failed to find. In their experiment, the younger children (age 3;10–5;2, mean: 4;7) accepted the test sentence only 35% of the time, while the adults in a control group accepted the statement 100% of the time. This result suggests that the younger children often failed to assign the inverse scope interpretation to the test sentence, leading to the low acceptance rate. Similarly, in another experiment, children ranging in age from 4;0 to 7;3 (mean: 5;11) accepted the inverse scope interpretation of the sentence in (6) only 7.5% of the time. The sentence was presented as a description of a situation where two horses jumped

---

**1** Lidz and Musolino (2002) extended these findings to Kannada, a language with SOV word order. They found that Kannada-speaking children have the same problem as English-speaking children in accessing wide-scope interpretations of object quantifiers, despite the difference in word order (negation follows the object in Kannada). Given this, Lidz and Musolino argued that children's scope interpretations are constrained by surface c-command relations between negation and quantifiers, not by linear word order.

over the fence, but a third one did not. As in the first experiment, the inverse scope interpretation (NOT >> ∀) makes the sentence true, whereas the surface scope interpretation (∀ >> NOT) is false. Children's justifications for their negative judgments also suggested that they were adhering to "isomorphic" scope interpretations, the interpretations that match surface word orders, and hence, this finding is called *Observation of Isomorphism*.

Musolino, Crain, and Thornton (2000) argued that children's adherence to isomorphic scope interpretations is derived from the application of the subset principle, a conservative learning algorithm. They proposed that there exists a binary parameter of UG, which distinguishes languages that only allow isomorphic scope interpretations from languages that allow more flexible scope interpretations. According to Musolino et al., Chinese is an example of the former type of language, where the counterpart of (6) permits only an isomorphic interpretation:

| | | | | | |
|---|---|---|---|---|---|
| (7) | *Mei-pi* | *ma* | *dou* | *mei tiao-guo* | *langan.* |
| | every-CLF | horse | all | not jump-over | fence |

'Every horse didn't jump over the fence.'
∀(x) [horse(x) → ¬ jump over the fence (x)] (every > not)
(Musolino, Crain and Thornton 2000: 22)

English, on the other hand, selects the other value of the parameter, hence permitting both isomorphic and inverse scope interpretations. Thus, the Chinese value of the parameter allows a subset of scope interpretations that are possible on the English value. In this scenario, first language learners must determine the correct parameter value for the target language. Musolino et al. claimed that the subset principle forces young children to choose the subset value to avoid the learnability problem associated with "unlearning" impossible scope interpretations. This initial setting of the parameter results in the non-adult adherence to isomorphic scope interpretations by young English-speaking children. Crucially, this approach to the observation of isomorphism assumes that children's non-adult behavior derives from their non-adult grammar: Young children adhere to isomorphic scope interpretation because that is the only possible option within their grammar.

This "grammatical" approach, however, has faced several difficulties. In more recent experimental works, it has been found that children's performance with inverse (i.e., non-isomorphic) scope interpretations is greatly improved by implementing certain changes to the context in which the experimental sentences are presented. For example, Gualmini (2003) found that children showed significantly less difficulty in accepting the inverse scope interpretation of sentences such as (5) when these negative sentences are used to indicate the discrepancy between a contextual expectation and the actual outcome. Musolino and Lidz (2002; 2006) also showed that children's performance related to inverse scope greatly improved when negative test sentences are preceded by a positive lead-in (e.g., *Every horse jumped*

*over the fence but every horse didn't jump over the barn*), possibly illustrating the same phenomenon. These new findings show that young children do have the ability to compute inverse scope interpretations, and strongly suggest that the original observation of isomorphism is derived from children's non-grammatical interpretive bias towards a certain type of scope interpretation. In other words, young children do not lack the grammatical device that inverts the scope of quantificational elements, and they are able to construct inverse scope representations, provided that the experimental context is properly controlled.

Note, however, that these new findings do not directly deny the possibility of the conservative learning of scope interpretations. The data only show that English-speaking children are able to construct scope interpretations that are available in their target language. Assuming that the input data for English-speaking children involves positive evidence for inverse scope interpretations in the relevant constructions (e.g., hearing "Every horse didn't jump over the fence" in a situation where some, but not all, horses jumped over the fence), it is possible that children have learned the availability of inverse scope through input data. In other words, the experimental data are still compatible with the position that the initial grammar is indeed restricted, as Musolino, Crain, and Thornton proposed. To determine whether some kind of conservative learning algorithm restricts the acquisition of scope, it is necessary to assess children's knowledge of a *constraint* on possible scope interpretations. If empirical evidence shows that children are not sensitive to a certain language-specific constraint on scope interpretation, and over-generate scope interpretations that are not possible in their target language, then such evidence conclusively shows that children do not rely on conservative learning in the acquisition of the scope constraint. In the next subsections, I will review such empirical findings.

## 3.2 Scope acquisition in Mandarin Chinese

The English universal quantifier *every* in the subject position may inversely take scope under negation. Thus, *every horse didn't jump over the fence* can mean that it was not the case that every horse jumped over the fence, implying that "only some horses did". In contrast to English, Mandarin Chinese has been argued to lack the inverse scope interpretation in the corresponding construction (see (7)). Given this cross-linguistic contrast, children acquiring Chinese must somehow learn the scope constraint. Under Musolino, Crain, and Thornton's conservative learning model, the relevant grammatical knowledge of Mandarin is a consequence of children's default hypothesis; children initially choose the "Chinese" value of the parameter due to the subset principle. Because adult Mandarin does not allow scope ambiguity with the relevant type of sentences, Chinese children would never encounter positive evidence for inverse scope. Consequently, if Chinese children are conservative learners

of scope, they will simply keep their initial hypothesis in the absence of falsifying evidence. In other words, the conservative learning approach predicts that young Chinese children show an adult-like scope interpretation pattern with sentences involving universal quantification and negation, disallowing inverse scope interpretations.

The study by Zhou and Crain (2009) provides an empirical test case for this prediction. They sought to determine whether Mandarin-speaking children are sensitive to the scope constraint. In one typical trial of their Truth Value Judgment Task experiment, they presented the following test sentence in a situation where every girl ate ice cream, but only one of them took pills:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (8) | *Mei-ge* | *nühai* | *dou* | *chi-le* | *bingjiling,* | *danshi* | *mei-ge* | *nühai* |
| | every-CLF | girl | all | eat-ASP | ice cream | but | every-CLF | girl |
| | *dou* | *meiyou* | *chi* | *yao.* | | | | |
| | all | not | eat | *pill* | | | | |

‘Every girl ate ice cream, but every girl didn’t take a pill.’

Since only some, but not every girl took pills, the test sentence is false under the surface scope interpretation (∀ >> NOT), but true under the inverse scope interpretation (NOT >> ∀). Thus, if a participant accepts the sentence as a valid description of the situation, then the response suggests that the participant has accessed the inverse scope interpretation. Note that the test sentence involves a positive lead-in before the crucial sentence with universal quantification and negation. This was necessary to exclude the possibility that some extra-grammatical factor prevents children from accessing particular scope interpretations. Remember that even English-speaking children in Musolino, Crain, and Thornton’s original experiment rejected the inverse scope interpretation of the test sentence in (6), while later experiments that included a positive lead-in greatly enhanced children’s performance (Musolino and Lidz 2002; 2006). Zhou and Crain also carefully constructed experimental stories so that the test sentences satisfied pragmatic felicity conditions for using negation.

The results are the following. First, none of the 20 Mandarin-speaking adults in a control group accepted the crucial test sentences (acceptance = 0%). This shows that the construction indeed disallows inverse scope interpretations in adult Mandarin. In contrast to this, 19 Mandarin-speaking children accepted the test sentence significantly more often (47%). Furthermore, the vast majority of inverse scope acceptance is from younger children. Of the 19 children, 9 younger children (ages 3;4–4;3) accepted the test sentences 89% of the time, while the acceptance rate of the 10 older children (ages 4;5–5;11) was only 10%. Finally, when the test sentence is presented in a situation where none of the girls take any pills (i.e., a situation that matches the surface scope interpretation of the test sentence), both adults and children showed 100% acceptance.

The results revealed that young Chinese children (around age 4) assigned non-adult inverse scope interpretations to sentences involving universal quantification and negation. This suggests that young children lack the knowledge of the language-specific scope constraint, hence allowing flexibility in syntax-semantics mapping. The empirical finding runs directly counter to the prediction of the conservative learning approach; Mandarin children do allow the scope interpretation that is not allowed in the target language, in addition to the adult-like surface scope interpretations. For the young Chinese children to become adults, they need to expunge their non-adult scope interpretations, and the experimental results suggest that they somehow accomplish this task before the age of six (i.e., older children correctly rejected the crucial test sentences). Since conservative learning does not provide an explanation for their acquisition of correct scope grammar, we will turn to an alternative learning model in Section 5.

## 3.3 Rigid scope constraint in Japanese

Goro (2007) carried out a set of experiments to determine whether Japanese- and English-speaking children access inverse scope interpretations in sentences such as (9) and (10):

| | | | |
|---|---|---|---|
| (9) | *Dareka ga* | *dono tabemono mo* | *tabe-ta.* |
| | someone NOM | every food | eat-PST |

(10) *Someone ate every food.*

The Japanese sentence (9) is scopally unambiguous for adult speakers, allowing only the interpretation that the subject existential quantifier takes scope over the object universal quantifier. In contrast, the English counterpart (10) allows the inverse scope interpretation, in addition to the surface scope interpretation. The surface scope interpretation of the sentence is false, for example, in a situation where each food was eaten by a different individual. The inverse scope interpretation of the sentence is true under the same situation. Thus, if the participant accepts the sentence in (9)/(10) presented in a situation where each of the foods was eaten by a different individual, the response suggests that the participant accessed the interpretation that makes the sentence true – the inverse scope interpretation. Conversely, if the participant does not allow the inverse scope interpretation, then she should reject the sentence in the same situation. The goal of the experiments is to compare the acceptance rates and patterns of Japanese- and English-speaking children to determine whether the difference in adult grammars affects children's behavior. Under a conservative learning model, it is predicted that Japanese children will adhere to the surface scope interpretation, disallowing the inverse scope interpretation. English children may or may not allow the inverse scope interpretation,

depending on the frequency with which they have encountered positive evidence for inverse scope interpretations.

The crucial test sentences were presented after a story about animals playing an "eating game". In the story, several groups of three animals of the same kind were invited to eat three pieces of food. In one condition, each animal in a group is generous and shares the snacks with his friends, making sure that every one of them gets to eat something. This pattern matches the inverse scope interpretation of the test sentences in (9) and (10): each food was eaten by a different individual. In the other condition, one of the animals in a group is greedy and eats all the food by himself. This pattern models the surface scope interpretation of the test sentences: a specific individual ate everything. The former condition provides the crucial test case for inverse scope, and the latter condition was included to ensure that the children do not have problems with the surface scope interpretation.[2] In fact, the Japanese children in a control experiment consistently accepted the test sentences under the latter condition (acceptance rate: 90.6%), showing no difficulties with the surface scope interpretation.

The crucial results from the main experiment are the acceptance rates of the inverse scope interpretation. Four different groups participated in the main experiment: (i) Japanese-speaking children (age 4;10–5;9, mean: 5;4); (ii) English-speaking children (age 5;0–5;10, mean: 5;4); (iii) Japanese-speaking adults; and (iv) English-speaking adults. To begin, let us review the adult behaviors. First, Japanese-speaking adults consistently rejected the inverse scope interpretation of the crucial test sentences (acceptance = 0%). This confirms the empirical claim that has been made in the theoretical literature – inverse scope interpretations are disallowed in Japanese. English-speaking adults, by contrast, accepted the inverse scope interpretation significantly more often than the Japanese-speaking adults, but the acceptance pattern is subject to inconsistency, and the overall acceptance rate was relatively low (33.6%). This acceptance rate resembles the data in previous studies on English-speaking adults (Kurtzman and MacDonald 1993; Marsden 2004), and the inconsistency probably reflects a general dispreference for inverse scope interpretations. Theoretical literature has acknowledged that inverse scope interpretations are subject to varying judgments, and adult speakers sometimes find it difficult to access an inverse scope interpretation when the corresponding surface scope interpretation is contextually plausible (e.g., Reinhart 2006). Given that the surface scope interpretation of the crucial test sentences was made plausible in the experimental story, some participants may have held on to that interpretation without considering another possible interpretation.

Turning now to the children, the acceptance rates for the inverse scope interpretation are the following: 42.2% for the Japanese-speaking children, and 35.9% for the

**2** Due to space limitations, the description of the experimental design is greatly simplified. See Goro (2007) for more accurate information.

English-speaking children. Interestingly, Japanese children accepted the crucial test sentences to a similar extent as the English-speaking children/adults, and significantly more often than the Japanese-speaking adults. This suggests that Japanese children share the underlying representation for inverse scope with English speakers, which is unexpected if children learn scope possibilities through conservative learning. The fact that Japanese-speaking adults never accepted the inverse scope interpretation of the test sentences shows that there is indeed a difference between the adult grammars of Japanese and English speakers, which presumably are reflected in the input data available to the children. Therefore, contrary to the prediction of a conservative learning model, the experimental findings suggest that Japanese children's grammar allows (over-)generation of particular scope interpretations without supporting evidence from their input. Similar "scope freedom" for children acquiring a scope-rigid language has been reported by several other studies (Chien and Wexler 1989; Sano 2004).

## 3.4 Scope reconstruction asymmetry in Japanese

It has often been pointed out that scrambled sentences in Japanese show a scope ambiguity that their canonical order counterparts lack (e.g., Hoji 1985). The following pair of canonical and scrambled order sentences illustrates the point:

(11) *Dareka ga* *daremo o* *semeta.*
someone-NOM everyone-ACC criticized
'Someone criticized everyone.'
$\exists >> \forall$ / $*\forall >> \exists$

(12) *Dareka* $o_i$ *daremo ga* $t_i$ *semeta.*
someone-ACC everyone-NOM criticized
Lit. 'Someone, everyone criticized.'
$\exists >> \forall$ / $^{OK}\forall >> \exists$

The scrambled sentence in (12) can be truthfully uttered in a situation where everyone criticized a different individual, and no single individual was criticized by everyone. Since the surface scope interpretation (i.e., there is a specific individual whom everyone criticized) should make the sentence false in this situation, this fact shows that the inverse scope interpretation is possible for the scrambled sentence, but not for its canonical-order counterpart in (11). A standard analysis of the contrast between (11) and (12) assumes that scrambling is a movement operation, and a moved phrase can be "reconstructed" to its base position at LF.

However, not all scrambled sentences show scope ambiguity. To illustrate, let us observe the interpretation of a sentence containing *dake* ‘only’ and *X mo Y mo* ‘both X and Y’:

(13) *Taroo dake ga* *huransugo* *mo* *supeingo* *mo* *hanasu.*
Taro-only-NOM French also Spanish also speak
‘Only Taro speaks both French and Spanish.’

The meaning of the sentence in (13) can be decomposed as follows:

(14) a. Taro speaks both French and Spanish, and
b. Nobody except Taro speaks both French and Spanish

In the meaning component (14b), the conjunction is interpreted under the scope of negation. Therefore, the sentence is true in the situation illustrated in (15) where Hanako speaks French but not Spanish, and Jiro speaks Spanish but not French.

(15)

| | Taro | Hanako | Jiro |
|---|---|---|---|
| French | ✓ | ✓ | * |
| Spanish | ✓ | * | ✓ |

However, the scrambled version of (14), as shown in (16), does not have the same interpretation. The sentence in (16) is false under the situation in (15):

(16) *Huransugo mo* *supeingo mo$_i$* *Taroo dake ga* $t_i$ *hanasu.*
French also Spanish also Taro-only-NOM speak
Lit. ‘Both French and Spanish, only Taro speaks.’

The interpretation of the sentence can be paraphrased as follows: With respect to French, Taro is the only one who speaks it; AND with respect to Spanish, Taro is the only one who speaks it. Here, the conjunction takes the widest scope, and therefore, the truth condition can be expressed by two conjoined propositions that each involve *only.* Crucially, the sentence lacks the scope interpretation that corresponds to the interpretation of the canonical-order sentence in (13). This indicates that scope-reconstruction with a scrambled conjunction is blocked by some independent constraint. The nature of this scope constraint will be discussed in Section 5.

Goro (2007) carried out a set of experiments that target this construction. The study sought to determine whether Japanese children permit the inverse scope interpretation of the sentence-initial NP with a conjunction, such as in sentences similar to (16). The experiment employs a standard truth value judgment task. The theme of the experimental story-line was a PSI-power demonstration, in which three cartoon characters (Pikachu, Doraemon, and Anpanman) attempted to perform various feats

using their PSI power (e.g., opening boxes without touching them, turning a frog into a princess, etc.) In one of the test trials, the three characters attempt to open two boxes, the blue and black boxes. In the story, Pikachu was the first one to attempt the opening of those boxes. He first opened the blue box successfully, and then the black box. Next, Doraemon made his attempt, but he failed to open the blue box. He moved on to the black box, but failed again. Anpanman was the last one, and he failed to open the blue box. Nevertheless, he did not give up and managed to open the black box. The final outcome of the story is illustrated in (17):

(17)

| | *Pikachu* | *Doraemon* | *Anpanman* |
|---|---|---|---|
| Blue box | ✓ | * | * |
| Black box | ✓ | * | ✓ |

At the end of the story, the puppet stated what he thought had happened during the trial, using the test sentence that involves a scrambled conjunction, as in (18):

(18) *Aoi hako mo kuroi hako mo*$_i$ *Pikatyuu dake ga* $t_i$ *aketa.*
blue box also black box also Pikachu-only NOM opened
Lit. 'Both the blue box and the black box, only Pikachu opened.'

Under the surface scope interpretation of the test sentence, the conjunction operator takes a wider scope than *dake* 'only', and the test sentence means that 'Only Pikachu opened the blue box, and only Pikachu opened the black box.' Under this interpretation, the sentence is false in the situation illustrated in (17), because Anpanman also opened the black box. In contrast, the inverse "reconstructed" interpretation makes the sentence true in the same situation. The inverse scope interpretation asserts that everyone other than Pikachu did not open both of the boxes, which is indeed the case in (17). The surface scope interpretation is the only interpretation that is acceptable for adults, and therefore adult speakers should reject the test sentence in this situation. Children should do the same if they obey the same restriction on scope interpretation as the adults.

The results demonstrate the following. First, adult Japanese speakers in the control group rarely accepted the crucial test sentences (acceptance rate = 7.8%), suggesting that the inverse scope/reconstructed interpretation is indeed unacceptable for adult Japanese speakers. In contrast, Japanese children (age: 4;11–5;10, mean: 5;6) were significantly more lenient about accepting the test sentences (acceptance rate = 76.6%). Therefore, once again, the experimental results provide evidence that children accessed a scope interpretation that was not allowed in their target language. Of the 16 children participating in the experiment, three acted similar to the adults in that they consistently rejected the test sentences. Those children must have somehow learned the constraint against scrambling reconstruction of which the remaining 13 children were unaware, a point we will discuss further in Section 5. Given that

children do over-generate reconstructed scope interpretations, the acquisition of the constraint cannot be a conservative learning process.

## 4 Freedom of scope and indirect negative evidence

The findings we reviewed so far point toward the same direction: young children allow scope interpretations that adults do not. In other words, children are *productive* in that they generate scope interpretations that are not included in the input data. Children's scope flexibility is observed across different languages, constructions, and combinations of quantificational elements. We refer to children's scope flexibility in general as *freedom of scope*. Freedom of scope has several implications about the nature of the learning mechanism that children use in the acquisition of possible scope interpretations. First, it strongly suggests that children's learning mechanism is not constrained by some general conservative learning algorithm. The scope constraints we have examined so far block a particular type of scope interpretation: inverse scope. Therefore, a single general conservative learning mechanism that blocks inverse scope in the absence of positive evidence can logically solve all learning problems in the acquisition of constraints. This is, however, not the solution that children adopt. Rather, they may productively assign a particular scope interpretation to a sentence without waiting for direct supporting evidence for that interpretation. Second, given that children allow scope interpretations that are impossible in the target language, the learning mechanism must involve some kind of non-conservative process that allows them to expunge their non-adult scope interpretations. In other words, a theory of language acquisition must now solve the difficult problem of "unlearning".

Seeking a solution to the problem, let us first consider the possibility that input evidence provides children with the necessary information for "unlearning." This amounts to asking whether children can extract any kind of negative evidence against a particular scope interpretation from the input they receive. In Section 1, I pointed out that it is extremely unlikely that children receive direct negative evidence in the form of parental feedback. Then, the remaining question is whether *indirect negative evidence* can provide children with the basis for expunging particular scope interpretations.

The idea of indirect negative evidence was first discussed in the early 1980s (e.g., Chomsky 1981) and has recently been attracting growing attention within the research on probabilistic learning models (e.g., Elman 1993; Lewis and Elman 2001; Seidenberg 1997; Tenenbaum and Griffiths 2001; Rhode and Plaut 1999; Regier and Gahl 2004). Roughly speaking, indirect negative evidence is the absence of input evidence that a certain hypothesis predicts to be possible in the language, and the learning mechanism uses the absence of expected data as evidence against the

hypothesis. An important characteristic of recent probabilistic learning models that shape learning around indirect negative evidence is that they have an ability to discriminate subset-superset hypotheses on the basis of positive evidence alone (e.g., Regier and Gahl 2004). Regarding the acquisition of possible scope interpretations, a probabilistic learner who detects the absence of a certain scope interpretation in the input data would be able to use this absence as evidence against the hypothesis that generates the scope interpretation.

As Pinker (1989: 15) points out, the use of indirect negative evidence is not strictly a feature of the input, but rather a feature of the child's learning mechanism. Therefore, the question of whether indirect negative evidence provides a plausible solution for the present problem needs to be tested with a particular model of a children's learning mechanism. To the best of my knowledge, no serious attempt has been made to construct a concrete probabilistic learning model for the acquisition of scope constraints. However, I have several reasons to doubt that probabilistic learning can solve the current problem of "unlearning".

One potential problem for a probabilistic learning scenario is that it requires an assumption that children can reliably identify the intended scope interpretations of input sentences. Since a probabilistic learning mechanism relies crucially on the absence of certain evidence, a probabilistic learner who is learning a ban against inverse scope must correctly detect the absence of the intended inverse scope in the input. However, the linguistic signals from input do not uniquely specify scope interpretations. For example, even if the learner hears the sentence *someone ate everything*, nothing about the form allows her to determine which of the two logically possible scope interpretations was intended. Rather, information about scope interpretations can only come from the learner's internally generated hypotheses about the meaning of the provided sentence. In other words, linguistic signals do not provide direct evidence for the existence and absence of a certain scope interpretation, and the discovery of such evidence depends on the learner's internal state.[3] Thus, given that children's grammar over-generates inverse scope interpretations at some point in development, it seems possible that their grammar wrongly assigns inverse scope interpretations to random input sentences that are uttered with intended surface scope. In fact, that seems to be exactly what children were doing in the experiments – they assigned to the test sentences an inverse scope interpretation

**3** I am aware of the fact that some marked intonation patterns can force a specific scope interpretation. However, such prosodic markings of scope do not seem to be a robust phenomenon. Leddon, Lidz and Pierrehumbert (2004) carried out an experiment in which they recorded English-speaking parents reading stories to their pre-school children. The stories contained potentially scope-ambiguous sentences such as *every bunny didn't jump over the fence*, and those sentences were read under two kinds of situations: ones that correspond to the surface scope interpretations and others that correspond to the inverse scope interpretations. The analysis of the recorded utterances found no systematic prosodic distinction between the intended surface scope and intended inverse scope interpretations.

that is not acceptable in the adult language. Such generation of non-adult scope can interfere with a probabilistic learning mechanism because it would lead to the "fabrication" of supporting evidence for impossible scope interpretations. In other words, even if input data lacks utterances with intended inverse scope, the learner's internal hypothesis may wrongly generate false evidence for inverse scope. This possibility can seriously undermine the necessary condition for a probabilistic learning mechanism to function correctly.

The second problem for a probabilistic learning scenario is that potentially informative input data can be very sparse. First, only sentences that involve two overt quantificational elements are relevant to the learners of possible scope interpretations. Second, among the possible combinations of quantificational elements, only a small subset of them is informative for learners. For example, *someone read a book* is not informative, because the surface and inverse scope interpretations are truth-conditionally indistinguishable. This inherent sparseness of relevant input information leads us to deduce that a probabilistic learning mechanism cannot extract any useful information from the input. For example, let us consider Japanese scrambled sentences. Remember that these sentences do not completely exclude inverse scope interpretations; with some combinations of quantifiers, reconstructed/inverse scope interpretation is possible. The relevant example is repeated here as (19):

(19) *Dareka* $o_i$ *daremo ga* $t_i$ *semeta.*
someone ACC everyone NOM criticized
Lit. 'Someone, everyone criticized.'
$\exists >> \forall$ / $\forall >> \exists$

Given this, children must learn to distinguish cases such as (19) from cases that do not allow an inverse scope interpretation, such as (20):

(20) *Huransugo mo supeingo* $mo_i$ *Taroo dake ga* $t_i$ *hanasu.*
Lit. 'Both French and Spanish, only Taro speaks.'
BOTH >> ONLY / *ONLY >> BOTH

For a probabilistic learning mechanism to discover the contrast between (19) and (20), the absence of inverse scope interpretations, as in (20), is not enough; it must be accompanied by substantial evidence for inverse scope interpretations, as seen in (19). Otherwise, a probabilistic learning mechanism would simply conclude that there is no difference between (19) and (20) with respect to possible scope interpretations. However, there are several reasons to suspect that such positive evidence for inverse scope interpretations is vanishingly rare. First, in colloquial Japanese, especially in child-directed speech, argument NPs are often dropped, making the vast majority of input sentences irrelevant to the learner regarding the relationship

between scrambling and scope. Second, Miyamoto and Nakamura's (2005) corpus study revealed that, in actual language use, scrambled word order is much less frequent than canonical word order. Thus, it is natural to assume that scrambled sentences with two (overt) quantificational arguments are accordingly rare, and more importantly, only a subset of those sentences is informative for learners of scope. Third, even if a child is fortunate enough to encounter such an exceptional example, evidence for inverse scope may only be obtained if the learner actually chooses to compute the interpretation, which may not always occur. These considerations naturally lead to an expectation that positive evidence for inverse scope, such as in (19), is very close to zero. If that is the case, then probabilistic learning models would face a serious challenge in discriminating between cases such as (19) and (20). The input data simply do not provide sufficient numbers of relevant cases that can be used for a probabilistic learning algorithm, and the two hypotheses are similar with respect to the amount of supporting evidence they receive from input. In other words, the inherent sparseness of positive evidence for inverse scope may trivialize the significance of absent evidence. With sparse data, positive evidence for grammatically possible scope interpretations can be equally absent as positive evidence for grammatically impossible scope interpretations, making it impossible to distinguish the two classes of scope interpretations on the basis of the absence of positive evidence.

In summary, given the specific properties of the relevant input and evidence received by the learner, the probabilistic learning of possible scope interpretations through indirect negative evidence appears to be quite implausible. This discussion is not meant to deny the general potential of probabilistic learning in all domains of grammar acquisition; probabilistic learning can play an important role in the acquisition of forms, such as word order and agreement. However, with respect to the acquisition of scope, an inherent indirectness and the sparseness of the relevant evidence pose particularly difficult problems for a probabilistic learning model. Given the absence of a concrete model that would solve the problems, I tentatively conclude that a theory of language acquisition must assume that children do not rely on negative evidence (direct or indirect) to correct their over-generating hypotheses regarding possible scope interpretations.

So far, I have indicated that children over-generate scope interpretations that are not acceptable in the target language, and therefore they must learn to purge their non-adult interpretations. I also argued that input data do not provide reliable negative evidence against children's non-adult interpretations, so children cannot rely on this to correct their grammar. Thus, we have now established the first two components of Pinker's learnability paradox. To construct an explanatory theory for the acquisition of scope, we must therefore challenge the third component: arbitrariness.

# 5 Towards a theory of non-arbitrary constraints

A constraint is arbitrary when its effect (e.g., the impossibility of a certain set of scope interpretations) cannot be derived from any other property of the grammar. Thus, learners of an arbitrary scope constraint must independently discover the impossibility of the relevant scope interpretations from input experience, and thus can be seriously impaired by data indirectness and sparseness. In contrast, if the impossibility of a certain set of scope interpretations is a consequence of some other property of the language, learners do not need to rely on input evidence to determine what is impossible. In this connection, an important insight can be drawn from the concept of *parameter* in the Principles and Parameters approach (Chomsky 1981, 1986).

When this concept was first introduced to the theory of grammar and language acquisition, parameter aimed to derive multiple consequences by setting the value of one parameter, thereby reducing the burden on an inductive learning mechanism. This original idea of parameter is clearly stated in the following quote from Chomsky (1981: 4):

> If these parameters are embedded in a theory of UG that is sufficiently rich in structure, then the languages that are determined by fixing their values one way or another will appear to be quite diverse, since the consequences of one set of choices may be very different from the consequences of another set; yet at the same time, limited evidence, just sufficient to fix the parameters of UG, will determine a grammar that may be very intricate and will in general lack grounding in experience in the sense of an inductive basis.

Over the years, the meaning of the term parameter has been stretched to the extent that it is sometimes used to refer to highly specific cross-linguistic contrasts. Nevertheless, in some recent parametric approaches to language acquisition (e.g., Snyder 2001, 2007; Sugisaki 2003), the original spirit of parameter remains the same; the system of parameters allows the learner to derive a wide variety of grammatical consequences by setting parameter values on the basis of limited evidence.

Quite independent of whether one actually employs some specific mechanism of parameters, the parametric approach to language acquisition provides an important insight into learners' acquisition of knowledge regarding unacceptable options. This insight can be stated in the following terms. For a learner who is equipped with a grammatical system that has a rich internal structure, learning something new (i.e., introducing a new component to the grammar) can affect parts of the existing grammar. Such a consequence can be negative in the sense that it may block the generation of representations that had previously been possible. Under this scenario, the learner acquires knowledge about unacceptable choices as a consequence of learning something new. This opens up the possibility of avoiding the data-sparseness problem that arises within the acquisition of scope. Suppose that a certain property

X in a language/construction is a consequence of another grammatical property Y in the language/construction. Then, as long as the learner knows the causal relationship between X and Y, and Y can be learned from observable properties in the input, then the learner does not need evidence regarding X from the input. The acquisition of X effectively piggybacks on the acquisition of Y.

Given the experimental data that show children do not obey certain constraints on scope interpretation, Zhou and Crain (2009) and Goro (2007) provide a theory that derives the effect of the constraints from other observable properties in the language. Due to space limitations, this chapter can only provide a rough sketch of the proposals; interested readers should consult the original papers for details. My aim here is not to examine the success of these specific proposals; rather, I would like to emphasize that the empirical data from language acquisition studies can create serious problems for a linguistic theory (i.e., a theory of adult grammar).

Let us begin with the unavailability of inverse scope in Mandarin sentences that include a universal quantifier and negation. A relevant example (21) follows:

| | | | | | | |
|---|---|---|---|---|---|---|
| (21) | *Mei-pi* | *ma* | *dou* | *meiyou* | *tiao-guo* | *liba.* |
| | every-CLF | horse | all | not-have | jump-over | fence |
| | $\forall >> \neg$ / $*\neg >> \forall$ | | | | | |

Zhou and Crain (2009) argue that the sentence involves a focus-sensitive operator, *dou*. Based on the assumption that *dou* is a focus operator, they propose that it induces cleft-like semantic structures. Thus, the logical form of the sentence in (21) corresponds to that of the English cleft construction in (22):

(22) *It was every horse that didn't jump over the fence.*

Under this analysis, the Mandarin sentence in (21) makes the claim that the focus element, 'every horse', has the property of not having jumped over the fence. Inverse scope is therefore impossible, given that the universal quantifier 'every' is focused. This analysis derives the impossibility of inverse scope from the focus-sensitivity of the lexical item *dou*. Consequently, if a child is not aware of the focus-sensitive property of *dou*, she is expected to be insensitive to the scope constraint in adult Mandarin, allowing inverse scope interpretations in sentences such as (21). For such a child, learning the focus-sensitive property of *dou* has the effect of purging previously possible inverse scope interpretations. This scenario obviates the need for negative evidence to make children expunge their non-adult inverse scope interpretations, as long as the focus-sensitive property of *dou* can be learned on the basis of available input. Zhou and Crain contend that the data attesting to the crucial property of *dou* is abundant in the input because it is often used as a focus operator in adult language.

Turning now to scope rigidity in Japanese, Goro (2007) argues that this is due to a semantic property in nominative subjects. It has been widely observed that nominative *ga*-marked subjects in Japanese exhibit a peculiar semantic characteristic (e.g., Kuroda, 1965; Kuno, 1973, among many others). The property is often referred to as the *exhaustive listing* implicature; sentences with a *ga*-marked subject imply that the subject represents an exhaustive list of entities that satisfy the predicate of the sentence in the relevant domain/context. Goro proposes that this semantic property of *ga*-marked subjects invokes the scope rigidity effect in Japanese canonical order sentences. For example, consider the following sentence (23) that exhibits scope rigidity:

(23) *Dareka ga* *dono* *tabemono mo* *tabe-ta.*
someone NOM every food eat-PST
'Someone ate every food.'
$\exists >> \forall$ / $*\forall >> \exists$

Since the subject is marked by *ga*, it carries an exhaustive listing implicature, and the sentence implies that 'someone' represents an exhaustive list of individuals who satisfy the predicate of the sentence. In other words, the semantic property of the *ga*-subject leads to an implicature that only one individual ate every food. However, this implicature is not compatible with the inverse scope/distributive interpretation of the sentence; the distributive interpretation entails that multiple individuals are referenced by the predicate (i.e., one eater per food). The exhaustive listing implicature thus blocks an inverse scope interpretation, yielding the scope rigidity effect in canonical order sentences. Under this approach, Japanese children are expected to allow inverse scope interpretations if they have not acquired the relevant semantic/pragmatic property of *ga*-marked subjects. Another possibility is that Japanese children at the age of 5 have already acquired the knowledge of *ga*-subjects, but they did not compute the pragmatic implicature in the experimental trials. It has been widely observed that young children do not reliably compute pragmatic implicature in a truth value judgment task (e.g., Noveck 2001; Papafragou and Musolino 2003; Guasti et al. 2005). In any case, the scenario predicts that once children start computing the implicature of *ga*-subjects, their non-adult inverse scope is blocked.

Let us now move on to the asymmetry of scope-reconstruction in scrambled arguments. First, following Bošković and Takahashi (2001), Goro (2007) assumes that "scrambled" phrases are base-generated in their surface positions. Those phrases may be LF-lowered to the "base" position to check their formal features. The crucial assumption here is that the LF-lowering is a last resort operation: it is allowed only when the scrambled phrase has some unchecked formal features that would otherwise lead the derivation to crash. Consequently, this approach predicts that a scrambled phrase that lacks unchecked formal features may not be LF-lowered, and

therefore, may not scope-reconstruct. An observation that is relevant to this model of scrambling is that *X mo Y mo* 'both X and Y' may never be followed by a case particle (nominative *ga*, accusative *o*, or dative *ni*), in contrast with other quantified NPs such as *dareka* 'someone':

(24) **Taroo mo Hanako mo ga / *Taroo mo Hanako mo o / *Taroo mo Hanako mo ni*

(25) *$^{OK}$dareka ga / $^{OK}$dareka o / $^{OK}$dareka ni*

Based on this observation, Goro argues that *X mo Y mo* lacks a formal case feature, and therefore the form cannot undergo the last-resort lowering. Consequently, a scrambled *X mo Y mo* is interpreted in its surface position, disallowing the inverse scope/reconstructed interpretation. Goro further proposes that children wrongly assign a case feature to *X mo Y mo*, and therefore they allow reconstruction (and consequently, inverse scope) for a scrambled *X mo Y mo*. Acquiring the proper feature composition of the expression *X mo Y mo* then has the effect of expunging non-adult inverse scope. Goro hypothesizes that once children notice that *X mo Y mo* is never followed by a case particle (via some kind of probabilistic analysis of the input data), they revise their default hypothesis that the noun phrase has a formal case feature.

Our discussion so far illustrates the impact of findings in child language research on a linguistic theory. The three proposals that I have reviewed in this section are theories of adult grammar as motivated by the empirical evidence revealed in studies on language development. The crucial observation is freedom of scope: children over-generate non-adult scope interpretations. Freedom of scope, combined with considerations on the (un)reliability of negative evidence, precludes theories of arbitrary scope constraints, however descriptively successful they are. In other words, whenever children generate a particular type of scope interpretation that is not acceptable in their target language, a grammatical theory must somehow derive the impossibility from some independently observable property of the language. Otherwise, the theory encounters the learnability paradox, and hence fails to attain explanatory adequacy. This straightforward relationship between acquisition data and a theory of adult grammar is mainly due to the nature of input evidence concerning scope interpretations. As I have discussed above, input evidence for scope is inherently indirect and very likely to be sparse, which leads to the conjecture that negative evidence (direct or indirect) does not play a direct role in the acquisition of this domain of knowledge.

# 6 Conservative learning of scope

So far, I have discussed cases that represent freedom of scope. Children do not obey language-specific constraints on scope and allow scope interpretations that are not permitted in the adult language. Freedom of scope strongly suggests that no general

conservative learning algorithm restricts all aspects of the acquisition of possible scope interpretations. It is possible, at least with particular constructions, for children to generate scope interpretations that are never exemplified in the input data. A question that arises here is to what extent children are productive/flexible with respect to scope assignment. This is an important question, because if children are maximally productive in assigning scope, the learnability consideration leads to a significant consequence: no language-specific constraint on scope interpretation can be arbitrary. Assuming that the considerations on the unreliability of negative evidence apply to other areas of scope acquisition as well, a productive learner who always allows every logically possible scope interpretation simply cannot learn any arbitrary constraint on scope. Conversely, if an arbitrary language-specific constraint on scope interpretation exists, children cannot be overly productive in the acquisition of such a constraint. In such a case, children must restrict their hypothesis concerning possible scope assignments so that they do not generate scope interpretations that cannot be expunged on the basis of positive evidence alone. Thus, whether children are uniformly productive in the acquisition of scope is an important empirical question that can have a huge impact on linguistic theory. In this section, I will explore empirical data that bear on this point. The data come from studies on the acquisition of logical connectives, where children show a distinctive pattern of behavior with respect to scope interpretation.

In English, when the disjunction operator *or* and the conjunction operator *both* … *and* … appear in the object position of simple negative sentences, they are interpreted within the scope of negation. Consequently, such sentences allow an inference that closely resembles De Morgan's laws of classical logic. To illustrate, the truth conditions of a sentence that contains a negated disjunction can be recast with the conjunction *and* presiding over both of the disjuncts, as shown in (26). With a negated conjunction, the truth conditions can be recast with the disjunction *or*, as in (27):

(26) *John doesn't speak Spanish or French.*
→ John doesn't speak Spanish <u>AND</u> doesn't speak French

(27) *John doesn't speak both Spanish and French.*
→ John doesn't speak Spanish <u>OR</u> doesn't speak French

In contrast, the Japanese counterparts of the sentences in (26) and (27) yield somewhat different interpretations. It appears that the disjunction *ka* and conjunction … *mo* … *mo* must take scope over local negation, and consequently, sentences with a negated logical connective do not allow De Morgan's inferences (Szabolcsi 2002; Goro and Akiba 2004; Goro 2007; Crain et al. 2006):

(27) *Zyon wa supeingo ka huransugo o hanasa-nai.*
John TOP Spanish or French-ACC speak-NEG
Lit. 'John doesn't speak Spanish or French.'
→ John doesn't speak Spanish OR he doesn't speak French.

(29) *Zyon wa supeingo mo huransugo mo hanasa-nai.*
John TOP Spanish also French also speak-NEG
Lit. 'John doesn't speak both Spanish and French.'
→ John doesn't speak Spanish AND doesn't speak French.

These interpretive contrasts between English and Japanese reveal the existence of another language-specific constraint on scope: Japanese logical connectives cannot take scope under local negation.

Goro and Akiba (2004) sought to determine whether Japanese children obey the scope constraint. In a typical trial of their truth value judgment task experiment, the participant was asked to judge whether sentence (30) or (31) was an accurate description of a situation in which the pig had eaten the carrot but not the green pepper:[4]

(30) *Butasan wa ninzin ka piman o tabe-nakat-ta.*
pig TOP carrot or pepper ACC eat-NEG-PST
Lit. 'The pig didn't eat the carrot or the pepper.'

(31) *Butasan wa ninzin mo piman mo tabe-nakat-ta.*
pig TOP carrot also pepper also eat-NEG-PST
Lit. 'The pig didn't eat both the carrot and the pepper.'

In adult Japanese, both the disjunction and the conjunction are interpreted as having scope over negation. The interpretation of (30) can be paraphrased as "The pig didn't eat the carrot, or he didn't eat the pepper", and therefore the sentence is true under this situation. In contrast, (31) means "The pig didn't eat the carrot, and he didn't eat the pepper", and is false under the same situation. As expected, adult Japanese speakers in a control group consistently accepted (30), and consistently rejected (31) (acceptance rate: 100% for (30), and 0% for (31)).

Children, however, showed different patterns of behavior. Of the 30 Japanese-speaking children (Age: 3;7–6;3, Mean: 5;3) participating, they accepted the test sentence with the disjunction (e.g., (30)) only 25% of the time, in sharp contrast to the adult's 100% acceptance rate. Furthermore, children's individual behaviors

4 Again, this is a vastly simplified description of the experimental design. The actual experiment involves several manipulations that are necessary to satisfy pragmatic felicity conditions for using negation and disjunction. See Goro and Akiba (2004) for details.

are highly consistent. Among the 30 children, only four followed adult response patterns in that they consistently accepted this type of test sentence. The acceptance rate of the remaining 26 children was 13%, i.e., they rejected the test sentences 87% of the time.

When the children who rejected the test sentence were asked to explain the reason for their negative judgments, most of them said either "because the pig did eat one of the vegetables" or "because it is only one of the vegetables that the pig didn't eat". The negative judgments of the vast majority of children, combined with their explanations for these, suggest that Japanese children interpreted the disjunction *ka* within the scope of negation, assigning the English-type "didn't eat the carrot and didn't eat the pepper" interpretation to (30). In contrast to this non-adult behavior with the disjunction, Japanese children almost unanimously agreed with the adult choice concerning the conjunction; they consistently rejected (31) when the pig ate the carrot but not the green pepper (acceptance rate = 5%). This behavior suggests that children assigned the adult-like wide-scope interpretation to the conjunction … *mo* … *mo*, rejecting the alternative narrow-scope interpretation that makes the sentence true.

In sum, Japanese children interpreted the disjunction *ka* as having scope under local negation, and did not allow the adult-like wide-scope interpretation. Conversely, children assigned the adult-like wide-scope interpretation to the conjunction … *mo* … *mo*, while correctly rejecting the interpretation in which the conjunction takes scope under negation. These patterns of behavior by Japanese children contrast with the children's performance in other studies we have reviewed. First, these behaviors are not due to a uniform bias towards a particular type of scope interpretation, such as isomorphic scope interpretation. Japanese children demonstrated different patterns of scope assignments toward the disjunction and the conjunction: narrow/isomorphic scope for a negated disjunction, and wide/inverse scope for a negated conjunction. Second, children restricted their scope interpretations to those particular patterns; i.e., they did not exhibit the scope flexibility observed in other studies. In the studies reviewed above, children *accepted* test sentences by assigning non-adult scope interpretations to them. However, in Goro and Akiba's experiment, children *rejected* the crucial test sentences (30) and (31), and this suggests that they did not allow the adult-like scope interpretation for a negated disjunction or the non-adult scope interpretation for a negated conjunction. Thus, children are not always "free" with respect to scope interpretation – they do not always allow a non-adult scope interpretation. The question then is, why did the children show this specific pattern of scope assignments with these sentences containing negation and logical connectives?

Goro (2007) proposed a parametric account for the acquisition of logical connectives (See also Crain, Goro, and Thornton 2006; Crain and Khlentzos 2008; Crain 2012). The background assumption is that the logical connectives in natural languages are divided into two classes. One class consists of logical connectives that are positive

polarity items (PPIs), which cannot be interpreted within the scope of local negation. This class includes the Japanese disjunction *ka* and the conjunction … *mo* … *mo*, along with the Hungarian disjunction *vagy* (Szabolcsi 2002), among others. The other class includes English and German logical connectives, which are not PPIs. Goro proposed that the polarity sensitivity of a logical connective is determined by setting the value of a binary parameter associated with each connective: [+PPI, −PPI]. Setting the value to [+PPI] renders a positive polarity to the lexical item, forcing it to take scope over local negation. Choosing the [−PPI] value results in a logical connective that has no such scope restriction, and hence yields De Morgan's interpretation in simple negative sentences. Under this account, the acquisition of logical connectives involves the discovery of the correct parameter value in the target language. This acquisition process, Goro argued, is restricted by the Semantic Subset Principle (Crain, Ni, and Conway 1994). The Semantic Subset Principle enforces an ordering of the values of certain parameters and compels children to adopt the value that yields the narrowest truth conditions as the default setting; this value is abandoned only on the basis of falsifying positive evidence in the input. Thus, the Semantic Subset Principle is a particular type of conservative learning algorithm. The crucial point here is that the Semantic Subset Principle establishes different default values for the [+PPI, −PPI] parameter for disjunction and conjunction: [−PPI] for disjunction and [+PPI] for conjunction. With respect to disjunction, the [−PPI] value of the parameter yields the "not A and not B" truth condition for a negated disjunction, as in English. A [+PPI] disjunction, however, yields the "not A or not B" truth conditions, as in Japanese. Notice that the two types of truth conditions stand in a subset/superset relationship; the situations in which "not A and not B" is true are a subset of the situations in which "not A or not B" is true. Therefore, according to the Semantic Subset Principle, the value that yields the narrower truth condition, namely the [−PPI] value, is selected as the default value for a disjunction.[5] With respect to conjunction, the relationship between relevant truth conditions is reversed: [−PPI] yields "not A or not B" and [+PPI] yields "not A and not B", with the latter truth condition being a subset of the former. Consequently, [+PPI] is the default value for a conjunction, because the value yields the subset truth condition. This parametric account, combined with the Semantic Subset Principle, predicts that children initially set the parameter value to [−PPI] for a disjunction and [+PPI] for a conjunction. Children, therefore, are expected to interpret a disjunction within the scope of local negation, and assign obligatory wide scope to a conjunction in simple negative sentences, irrespective of the properties in their target

---

**5** This argument presupposes that a [−PPI] disjunction may not take scope over negation, disallowing scope ambiguity. Descriptively, this seems to be the case; English-speakers consistently reject the wide-scope interpretation of disjunction in sentences like *John doesn't speak Spanish or French*. The underlying mechanism that blocks the wide-scope disjunction is still unclear. See Goro (2007) and Jing (2008) for discussion.

language. Japanese children's behaviors in Goro and Akiba's experiment bear out this prediction.

Several pieces of supporting evidence for the parametric account have been found in various languages. For example, Komine (2012) replicated Goro and Akiba's results with Japanese simple negative sentences and with comparative constructions. Jing, Crain, and Hsu (2005) and Verbuk (2006) found that Chinese and Russian children, respectively, rejected the wide-scope interpretation of disjunction over negation, even though the interpretation is readily accepted in those (adult) languages. With respect to conjunction, Crain et al. (2013) found that English- and Chinese-speaking children did not assign narrow-scope interpretations to a conjunction in simple negative sentences. For example, English-speaking children almost never accepted the sentence *The pig didn't eat both the pepper and the carrot*, when the pig had eaten the carrot but not the pepper. In short, across languages (Japanese, Chinese, Russian, and English), children's interpretations of the scope of logical connectives in negative sentences show exactly the same pattern: disjunction is interpreted under the scope of negation, and conjunction is interpreted over the scope of negation. Furthermore, children are equally conservative in scope assignment; they rejected the alternative scope interpretations, even if the scope assignment makes the test sentence true.

It is important to observe that the domain to which the Semantic Subset Principle is applied seems to be restricted. To illustrate, let us review Goro and Akiba's (2004) control experiment, where they replaced the disjunctive phrase in the test sentence with an indefinite existential, *nanika* (something):

| | | | | |
|---|---|---|---|---|
| (32) | *Butasan* | *wa* | *nanika* | *tabe-nakat-ta.* |
| | pig | TOP | something | eat-NEG-PST |
| | Lit. 'The pig didn't eat something.' | | | |

Here, just as in the case with disjunction, the narrow scope of *nanika* (i.e., NOT >> ∃: The pig didn't eat anything) yields a subset truth condition of the wide-scope interpretation (i.e., ∃ >> NOT: There is something that the pig didn't eat). Therefore, if the Semantic Subset Principle restricts every aspect of children's scope assignment in the same way, then Japanese children should show a bias towards the subset interpretation, as they did with the negated disjunction. Japanese children, however, did not show such a bias. The test sentence was presented in a situation where the pig had eaten, for example, the carrot and the eggplant, but not the pepper. Children accepted the sentence 88% of the time, showing that they accessed the wide-scope interpretation of *nanika*. A similar observation can be found in Musolino, Crain, and Thornton (2000), where they assessed English-speaking children's interpretation of negative sentences with *every* in the object position:

(33) *The Smurf didn't buy every orange.*

The experimental story was that the Smurf examined three oranges, and ended up buying only one of them. Children accepted (33) as a correct description of the story 85% of the time. This behavior suggests that children accessed the narrow-scope interpretation of *every*, because the alternative wide-scope interpretation of *every* would make the sentence false. This result contrasts with the observation in Crain et al. (2013), where English-speaking children almost never allowed the narrow-scope interpretation of conjunction (*both X and Y*) in the object position of simple negative sentences. Unlike conjunction, the universal quantifier *every* does not elicit children's bias towards a subset truth condition. Zhou and Crain's (2009) Chinese data (discussed above) leads to the same conclusion. In their experiment, young Mandarin-speaking children interpreted a universal quantifier under the scope of negation, showing no adherence to the surface scope that yields a subset truth condition.

The Semantic Subset Principle is a conservative learning algorithm; it forces the learner to choose the narrowest possible hypothesis, and thereby, prohibits other possibilities until the learner encounters a sufficient amount of falsifying positive evidence. The available evidence suggests that children utilize this type of conservative learning mechanism in the acquisition of relative scope between negation and logical connectives. This is expected if polarity sensitivity (in our terms, +PPI or −PPI) is an arbitrary lexical property of logical connectives in natural languages. If the property cannot be discovered by observing other independent properties, then conservative learning is the only way for children to acquire the scope constraint, assuming that negative evidence is unreliable. At the same time, the effects of the conservative learning algorithm seem to be restricted to this specific domain. Combined with the observation of scope freedom in other areas, this restricted conservatism suggests that children take advantage of multiple learning strategies in the acquisition of scope. Children are neither uniformly productive nor conservative with respect to scope assignment. The domain of knowledge seems to be partitioned into several distinct areas according to the nature of relevant quantificational elements, and children apply different learning strategies to each of the areas.

# 7 Conclusion

In this chapter, I have examined empirical evidence concerning the acquisition of language-specific constraints on scope interpretation in terms of learnability. A theory of language acquisition faces a learnability paradox if an acquisition task has the following three characteristics: (i) *productivity*; (ii) *no negative evidence*; and (iii) *arbitrariness*. Any adequate explanatory theory must therefore deny at least one of the three components. With respect to the acquisition of constraints on scope, the inherent properties of relevant evidence in the input make the second component,

*no negative evidence*, especially difficult to refute. Evidence concerning scope interpretation is inherently indirect, and tends to be very sparse. These properties lead to difficulties in constructing a realistic learning model that takes advantage of indirect negative evidence. These considerations leave us with two logical possibilities: either children are not productive or the relevant constraint on scope is not arbitrary. We have reviewed evidence that both of these possibilities are embodied in the actual acquisition process. First, with some constructions and combinations of quantificational elements, children exhibit freedom of scope. Japanese children, for example, allow non-adult inverse scope interpretations for sentences that involve two quantificational arguments. Given that children are undeniably productive in those areas, we are forced to conclude that the relevant scope constraints are not arbitrary; a linguistic theory must somehow derive the effects of the constraint from the interactions of other independently observable properties of the language. Second, with respect to the scope of logical connectives in negative sentences, children are highly conservative. Across different languages, children restrict their scope interpretations to those that yield subset truth conditions. These observations open up the possibility that human languages involve some arbitrary language-specific constraints on scope, as long as they are associated with a conservative learning algorithm.

One important consequence that we can draw from these empirical observations is that a "general-purpose learning mechanism" cannot explain the acquisition of scope. Rather than resorting to a single learning mechanism that would produce similar developmental patterns across different areas of scope acquisition, children employ different learning strategies according to the nature of the target constraint. The effect of a specific learning strategy is not carried over to highly similar areas (e.g., no effects of the Semantic Subset Principle in the acquisition of relative scope between negation and universal quantifier), so it seems that "choice" among the learning strategies is determined by some innate mechanism. In other words, there must be some innate mechanism that constrains a particular learning strategy from being applied to a highly restricted domain. Additional empirical evidence about children's acquisition of a particular constraint on scope will thus broaden our understanding about the nature of the learning principles and the constraints on them. This is an important issue for future research.

# References


Aoun, Joseph and Audrey Li. 1989. Constituency and scope. *Linguistic Inquiry* 20. 141–172.

Baker, C. L. and John. J. McCarthy. 1981. *The logical problem of language acquisition.* Cambridge, MA: MIT Press.

Beck, Sigrid and Shin-Sook Kim. 1997. On *wh-* and operator scope in Korean. *Journal of East Asian Linguistics* 6. 339–384.

Berwick, Robert. 1985. *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.

Bošković, Željko and Daiko Takahashi. 1998. Scrambling and last resort. *Linguistic Inquiry* 29. 347–366.

Brown, Roger and Camille Hanlon. 1970. Derivational complexity and order of acquisition in child speech, 11–54. In John R. Hayes (ed.), *Cognition and the development of language*. New York, NY: Wiley.

Chien, Yu-Chin and Kenneth Wexler. 1989. Children's knowledge of relative scope in Chinese. *Papers and Report in Child Language Development* 28. 72–80.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. Westport, CT: Praeger.

Clark, Robin. 1992. The selection of syntactic knowledge. *Language Acquisition* 2. 83–149.

Crain, Stephen. 2012. *The Emergence of meaning*. Cambridge: Cambridge University Press.

Crain, Stephen, Takuya Goro, Anna Notley, and Peng Zhou. 2013. A parametric account of scope in child language. In Stavroula Stavrakaki, Martina Lalioti and Polyxeni Konstantinopoulou (eds.), Advances in Language Acquisition, 63–71. Cambridge Scholars Publishing.

Crain, Stephen, Takuya Goro, and Rosalind Thornton, R. 2006. Language acquisition is language change. *Journal or Psycholinguistic Research* 35. 31–49.

Crain, Stephen, Weijina Ni, and Laura Conway. 1994. Learning, parsing, and modularity. In Charles Clifton, Jr., Lyn Frazier and Keith Rayner (eds.), *Perspectives on sentence processing*, 443–466. Amsterdam: Lawrence Erlbaum.

Crain, Stephen and Drew Khlentzos. 2008. Is logic innate? *Biolinguistics* 2(1). 24–56.

Elman, Jeffrey L. 1993. Learning and development in neural networks: The importance of starting small. *Cognition* 48. 71–99.

Fodor, Janet D. 1992. Learnability of phrase structure grammars. In Robert Levine (ed.), *Formal grammar: Theory and implementation* (Vancouver Studies in Cognitive Science 2), 3–68. Oxford, U.K.: Oxford University Press.

Fodor, Janet D. 1994. How to obey the subset principle: binding and locality. In Barbara Lust, Gabriella Hermon, and Jaklin Kornfilt (eds.), *Syntactic theory and first language acquisition: Cross-linguistic perspectives (vol. 2): Binding, dependencies, and learnability*, 429–451. Hillsdale, NJ: Lawrence Erlbaum.

Goro, Takuya. 2007. *Language-specific constraints on scope interpretation in first language acquisition*. University of Maryland at College Park dissertation.

Goro, Takuya and Sachie Akiba. 2004. The acquisition of disjunction and positive polarity in Japanese. In Vinetta Chand, Ann Kelleher, Angelo J. Rodriguez, and Benjamin Schmeiser (eds.), *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, 251–264. Somerville, MA: Cascadilla Press.

Gualmini, Andrea. 2003. Some knowledge children don't lack. In Barbara Beachley, Amanda Brown, and Frances Conlin (eds.), *Proceedings of the 27th Boston University Conference on Language Development*, 276–287. Somerville, MA: Cascadilla Press.

Guasti, Maria Teresa, Gennaro Chierchia, Stephen Crain, Francesca Foppolo, Andrea Gualmini and Luisa Meroni. 2005. Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes* 20. 667–696.

Hoji, Hajime. 1985. *Logical form constraints and configurational structures in Japanese*. University of Washington dissertation.

Huang, James C.-T. 1982. *Logical relations in Chinese and the theory of grammar*. MIT Dissertation.

Jing, Chunyuan. 2008. *Pragmatic computation in language acquisition: Evidence from disjunction and conjunction in negative context*. University of Maryland at College Park dissertation.

Jing, Chunyuan, Stephen Crain and Cheng-Fen Hsu. 2005. The interpretation of focus in Chinese: child vs. adult language. In Yukio Otsu (ed.), *Proceedings of the Sixth Tokyo Conference on Psycholinguistics*, 165–190. Tokyo: Hituzi Syobo.

Kim, Soo-Won. 1989. The QP status of wh-phrases in Korean and Japanese. In E. Jane Fee (ed.) *Proceedings of the Eighth West Coast Conference on Formal Linguistics*, 358–372. Stanford: CSLI Publications.

Komine, Emi. 2012. Japanese children's interpretation of a disjunction in comparative sentences and negative sentences. In Yukio Otsu (ed.), *Proceedings of the Thirteenth Tokyo Conference on Psycholinguistics*, 103–119. Tokyo: Hituzi Syobo.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kuroda, S.-Y. 1965 *Generative grammatical studies in the Japanese language.* MIT dissertation.

Kurtzman, Howard S. and Maryellen C. MacDonald. 1993. Resolution of quantifier scope ambiguities. *Cognition* 48. 243–279.

Leddon, Eric. M., Jeffrey Lidz and Janet Pierrehumbert. 2004. Suprasegmental cues to meaning in child-directed speech. Paper presented at the 2004 CUNY Conference on Human Sentence Processing. 25–27 March.

Lewis, John D. and Jeffrey L. Elman. 2001. Learnability and the statistical structure of language: Poverty of stimulus arguments revisited. In Barbora Skarabela, Sarah Fish, and Anna H.-J. Do (eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development*, 359–370. Somerville, MA: Cascadilla Press.

Lidz, Jeffrey and Julien Musolino. 2002. Children's command of quantification. *Cognition* 84. 113–154.

Manzini, Rita M. and Kenneth Wexler. 1987. Parameters, binding theory, and learnability. *Linguistic Inquiry* 18. 413–444.

Marsden, Heather. 2004. *Quantifier scope in no-native Japanese: A comparative study of Chinese, English, and Korean-speaking learners.* University of Durham dissertation.

Miyamoto, Edson and Michiko Nakamura. 2005. Unscrambling some misconceptions: A comment on Koizumi and Tamaoka (2004). *Gengo Kenkyu* 128. 113–129.

Musolino, Julien, Stephen Crain and Rosalind Thornton. 2000. Navigating negative quantificational space. *Linguistics* 38(1). 1–32.

Musolino, Julien and Jeffrey Lidz. 2002. Preschool logic: truth and felicity in the acquisition of quantification. In Barbora Skarabela, Sarah Fish, and Anna H.-J. Do (eds.), *Proceedings of the 26th Boston University Conference on Language Development*, 406–416. Somerville, MA: Cascadilla Press.

Musolino, Julien and Jeffrey Lidz. 2006. Why children aren't universally successful with quantification. *Linguistics* 44. 817–852.

Noveck, Ira. 2001. When children are more logical than adults: Experimental investigations on scalar implicatures. *Cognition* 78. 165–188.

Papafragou, Anna and Julien Musolino. 2003. Scalar implicatures at the semantics-pragmatics interface. *Cognition* 80. 253–282.

Pinker, Steven. 1979. Formal models of language learning. *Cognition* 7. 217–283.

Pinker, Steven. 1989. *Learnability and cognition.* Cambridge, MA: MIT Press.

Regier, Terry and Sussane Gahl. 2004. Learning the unlearnable: The role of missing evidence. *Cognition* 93. 147–155.

Reinhart, Tanya. 2006. *Interface strategies*. Cambridge, MA: MIT Press.

Rohde, Douglas and David Plaut. 1999. Language acquisition in the absence of explicit negative evidence: How important is starting small? *Cognition* 72. 67–109.

Roeper, Thomas and Jill de Villiers. 1992. Ordered decisions in the acquisition of wh-questions. In Jürgen Weissenborn, Helen Goodluck and Thomas Roeper (eds.), *Theoretical Issues in language acquisition: Continuity and change in development*, 191–236. Hillsdale, NJ: Lawrence Erlbaum.

Sano, Tetsuya. 2004. Scope relations of QP's and scrambling in the acquisition of Japanese. In Jacqueline van Kampen and Sergio Baauw (eds.) *Proceedings of the GALA 2003 conference on language acquisition*, 421–431. Netherlands: Netherlands Graduate School of Linguistics.

Seidenberg, Mark. 1997. Language acquisition and use: Learning and applying probabilistic constraints. *Science* 275. 1599–1603.
Snyder, William. 2001. On the nature of syntactic variation: Evidence from complex predicates and complex word-formation. *Language* 77. 324–342.
Snyder, William. 2007. *Child language: The parametric approach*. Oxford, U.K.: Oxford University Press.
Sugisaki, Koji. 2003. *Innate constraints on language variation: Evidence from child language*. Storrs, Connecticut: University of Connecticut dissertation.
Szabolcsi, Anna. 2002. Hungarian disjunctions and positive polarity. In István Kenesei and Péter Siptár (eds.), *Approaches to Hungarian* 8. 217–241. Budapest: Akademiai Kiado.
Tenenbaum, Joshua and Thomas Griffiths. 2001. Generalization, similarity and Bayesian inference. *Behavioral and Brain Sciences* 24. 629–640.
Verbuk, Anna. 2006. The acquisition of the Russian *or*. Paper presented at Western Conference on Linguistics (WECOL 06), California State University at Fresno, 27–29 October.
Wexler, Kenneth. 1993. The subset principle is an intensional principle. In Eric J. Reuland, Werner Abraham, and F. R. Ankersmit (eds.), *Knowledge and language: From Owell's problem to Plato's problem* (Vol. 1), 217–239. Dordrecht: Kluwer.
Wexler, Kenneth and Peter W. Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.
Zhou, Peng and Stephen Crain. 2009. Scope assignment in child language: Evidence from the acquisition of Chinese. *Lingua* 119. 973–988.

Masahiko Minami

# 6 Narrative development in L1 Japanese

## 1 Introduction

Narrative is defined as a form of extended discourse in which at least two different events are described so that a variety of relationships between them, such as temporal, causal, and contrastive, become explicit. All of us have stories to tell. Wherever we go, we find that narratives are used in such important functions as mediating interpersonal relationships, self-presentation, making sense of experiences, and, when it comes to the development of children's language skills, narration is the platform for transition into literacy. It is certainly true that narratives are typical of human discourse activity. To communicate with people around them, individuals not only of all ages but also of all cultural backgrounds need the ability to narrate. Parallel with this universality, however, a starkly culture-specific narrative style undeniably exists; the ways in which individuals talk about past events are deeply cultural.

The most fundamental feature of language development after the age of five is the change in function of linguistic categories from the sentential level to the level of extended spans of narrative discourse (Dickinson and McCabe 1991; Karmiloff-Smith 1986). The process that links early emergence of language development and relatively late mastery of linguistic knowledge represented by narrative discourse, in fact, begins much earlier. At around 22 months, children start to refer to past events, initially with a great amount of adult assistance. Around two years of age, children begin to narrate past events[1] such as injuries and to evaluate such experiences. Between three and five years old, children increase the length and complexity of their personal narratives. Actually, while English-speaking children tell various types of narratives (e.g., personal anecdotes and fantasies), the majority of their conversational narratives relate to real personal experiences (e.g., Miller and Sperry 1988; Preece 1987). We may therefore claim that the ability to tell a coherent personal narrative is of significant importance in many aspects of an individual's life, irrespective of age.

Furthermore, beginning at birth, children are immersed in a particular type of language and culture, the language and culture of the immediate community, which

1 This means that children begin to narrate memories at around two years of age, and they then proceed to develop their narrative abilities over the next few years. For instance, a two-year-old child may say "Big mean ugly fish" instead of "It was (or I saw) a big mean ugly fish," but this type of narrating memories takes place in conversational exchanges. Dickinson and McCabe (1991: 7) and McCabe (1993: 292) use the following example: A mother asks her 31-month-old son, "Did you like the puppy?" He replies, "He taste my knee" (without the inflection of the verb, i.e., instead of "He tasted my knee"). She echoes her son's reply but recasts it with a rising intonation, "He tasted your knee?" Her son replies, "Yeah. An' puppy chase me" (instead of "The puppy chased me").

includes the home and the school. The ability to narrate capably in a language- and culture-specific way is critical to later achievements in various developmental domains such as literacy. Despite the importance of understanding children's narrative development, a criticism of this field is that there is a dearth of work considering linguistic and cultural variations (e.g., Au 1993; Heath 1983; Michaels 1991; Philips 1982).[2]

To address this gap, this chapter explores how language shapes and is shaped by culturally specific experiences through analysis of: (I) how young children develop narrative structure, and (II) how adults/parents guide their children in the acquisition of culturally appropriate styles of narrative and literacy. Specifically, the first section of this chapter presents an overview of some theoretical approaches to language development in general, and narrative development in particular. The second section, which emphasizes sociocultural aspects, examines children's narrative discourse styles with particular attention to age related commonalities and differences. The section then discusses the role of parental input in facilitating the development of children's personal narratives, based on the belief that the origins of narrative style can be traced back to early conversations at home between parents and children. The third and final section provides an overall summary and discusses the relationship between sociocultural background and the development of literacy in young children. In this way, each section provides a framework for consideration of some important issues regarding sociocultural contexts in narrative discourse and emergent literacy.[3]

As the review of past research progresses, the chapter increasingly makes explicit the author's position as a social interactionist who believes that the child is not only under the influence of the environment but simultaneously acts upon and even creates the environment to a certain extent. This assumption is based on the social interaction paradigm advocated by Vygotsky (1978) and Bruner (1977). The chapter also makes clear the author's position as a new environmentalist in emphasizing the critical issue of cross-cultural differences, in particular the substantial cross-cultural differences in the ways in which children structure their narratives.

**2** Philips (1982), for example, describes how, because of differences in unconscious interactional norms, the verbal as well as nonverbal communicative style of Native American students causes conflicts and misunderstandings in interactions with European American teachers.

**3** In this chapter, I adopt Sulzby and Zecker's (1991: 195) definition of emergent literacy, young children's "everyday encounters with the print in their environments". Here, however, I define literacy as both spoken and written language. This is a much broader definition of literacy than is generally accepted, but I justify it by my belief that a strong connection exists between learning to talk (particularly learning to narrate) and learning to read and write. In other words, oral language development (narrative development in particular) is directly related to the later development of written language.

## 1.1 Differing views of language development

Science in different disciplines operates on different sets of assumptions. Unlike natural science, which is interested in the structure of natural phenomena, social science finds its interest in the normative grounds of individuals' actions, beliefs, goals, norms, rules, and values. In the past, researchers had been interested in the question of the compatibility of different forms of research such as "positivist versus constructivist" and "empiricist versus nativist". For example, the positivist philosophers of science, as compatibilists, assumed that all phenomena could be explained deductively, and they sought a unity of science. In contrast, incompatibilists, as the critics of positivism, did not believe that human actions could be deductively explained, and, instead, emphasized the necessity of interpretive descriptions and explanations with regard to the actor's intentions, motives, and purposes. With these diverse scientific disciplines as backdrops, reviewing past theories and considering their implications is an interesting endeavor when considering narrative development.

## 1.2 Behavioristic approach vs. linguistic approach

Positivism, which refers to a set of epistemological perspectives and philosophies of science, holds that the scientific method is the best approach to uncovering the processes by which both physical and human events take place. This was true of explanations about child development, most of which were dominated by behaviorist interpretations (Watson 1913, 1924). The behaviorist dominance was especially true in Western societies in general, and the United States in particular. Behaviorists, also called learning theorists or environmentalists, viewed the environment as molding the child. According to them, learning changes the child's behavior and advances his or her development.

While the behavioristic approach claimed that the learning principle of reinforcement plays the major role in the process of language acquisition, Chomsky (1959) argued that because the linguistic environment cannot account for the structures that appear in children's language, aspects of language rules and structure must be innate. He specifically criticized the behaviorist approach represented by Skinner (1957) who claimed that language learning is based upon experience. Instead, according to Chomsky (1957, 1965), humans have a biological or "hard-wired" endowment, because patterns governing natural languages are similar across different languages, though variations exist, and these variations are not always acquirable by children with the available input. The environment plays a role in the maturation of language (including language variation) and lexical learning. Thus, the behavioristic approach viewed children as beneficiaries of the language training mainly employed by their parents, whereas the generative approach represented by Chomsky considered children to be endowed with specialized language processors.

## 1.3 Cognitive-interactionist approach vs. social interaction approach as sociocognitive theory

Unlike behavioristic or opposing linguistic approaches, Piaget (1959 [1926]) argued that the complex structures of language seem neither innate nor learned. Rather, based on constructivism, Piaget's interactive approach suggests that language structures emerge as a result of continuing interaction between the child's current level of cognitive functioning and his or her linguistic as well as nonlinguistic environment.

On the other hand, the social interaction approach (or the sociocognitive theory) represented by Vygotsky (1978) combines the two opposing approaches of the behavioristic and the linguistic, and, furthermore, considers the functions of language in social communication to hold significant meaning throughout development. Vygotsky (1962) argued that language is at first only a tool for social interaction, but that the role of language changes over the course of development from a social tool to a private tool, as the child internalizes linguistic forms.

The above conceptualization of language development forms a basis for the view of language as a socioculturally mediated product. More generally, the conceptualization rests on the "constructivist" conception of meaning, stipulating that social interactions are culturally constrained. Vygotsky (1978) hypothesized that children learn from other people, and particularly that children's problem-solving skills, which include language, first develop through social interactions with more capable members of society – adults and peers – and then become internalized after long practice. Vygotsky's concept of the "zone of proximal development", which clarifies the difference between what learners can do without help and what they can do with help, is a construct that helps elucidate how interaction contributes to children's development. Through the process of social interaction, adults provide children with the tools to establish complex series of actions in problem-solving situations. Because the process of social interaction occurs before children have the mental capacities to take appropriate actions to solve problems on their own, adults need to regulate children's actions. Then, these regulatory behaviors taken by adults gradually become part of the children's own behavior. Unlike Skinnerian conditioning (Skinner 1957), the relationship between cognitive development and social interaction, particularly early social interactions between children and more mature members of society, can be summarized by Vygotsky's claim that all higher mental functions appear twice in development: (1) first as social or interpsychological functions during interactions with other social agents, and (2) only later, through the internalization of social-interactive processes, as individualized or intrapsychological functions.

When focusing on language development, like Piagetian theory (e.g., Piaget 1959), sociocognitive theory regards language development as part of more general cognitive development, emphasizing the acquisition of higher-order intellectual skills (e.g., Vygotsky 1978). In contrast to Piagetian theory, however, the central tenet

of sociocognitive theory is based on the foundational role that social interaction plays in cognitive and language development. In this sense, sociocognitive theory is highly compatible with language socialization studies (e.g., Schieffelin and Ochs 1986). The social interaction approach thus combines many aspects of both the behaviorist positions and the linguistic positions. While social interactionists agree with linguists that language has structure and follows certain rules (which make language unique or different from other behaviors), they also agree with behaviorists in terms of the role of the environment; i.e., that the structure of human language arises from social and communicative functions that language plays in human relations.

## 1.4 Evaluation of the various approaches reviewed

As seen in the claims made by both cognitive interactionists and social interactionists, examining the development of children's pragmatic ability is critical in order to understand language development. Chomsky's linguistic revolution (Chomsky 1957), however, brought with it the importance of conceptualizing links between the role of language and the human mind (i.e., humans have some universal and innate ability to learn language). The corrective emphasis on biology, however, is an oversimplification just as extreme as the Skinnerian one, though on the opposite end of the nature-nurture continuum. That is, in addition to the inborn capacity to analyze the underlying rule-governed nature of the target language, serious consideration should be paid to early socialization in order to explain the linguistic competence that young children acquire and develop. As Snow and Ferguson (1977) suggest, it seems obvious that parents play a far more important role in their children's language acquisition than simply modeling the language and providing input for what Chomsky (1965) claims is "a language acquisition device".

Those opposing paradigms for understanding language acquisition insisted that environmental influence on language development not be overlooked. The emergentist approach (MacWhinney 1999), which, in fact, adopted the aforementioned social interaction approach, claimed that domain-general cognitive mechanisms, such as working memory, statistical learning, and pressures on memory organization and retrieval, contribute to language acquisition, although an innate, domain-specific mechanism might allow the very initial emergence of language. More specifically, in the emergentist framework, domain-general cognitive mechanisms work on environmental stimuli to render the complex and elegant structures that characterize language. Thus, the emergentist view is a constructivist one that emphasizes the interaction between the organism and the environment (i.e., children gradually learn through interacting with environmental factors such as parents' speech patterns).

### 1.5 A sociolinguistic account: From language to narrative and to literacy development

Following the sociolinguists' account, language can be considered as both a manifestation and product of a culture and possibly, a social class. Consequently, people in different cultures have differing beliefs about how children learn language. Telling a story, in fact, involves many factors, some of which add diverse cultural flavors to a narrative. For example, people from different linguistic backgrounds (and possibly cultural backgrounds as well) might encode their own perspectives and emotions in distinct ways. Au (1993: 113) describes "talk story", an important speech event for Hawaiian children in their local speech communities: "During talk story children present rambling narratives about their personal experiences, usually enhanced with humor, jokes, and teasing. The main characteristic of talk story is joint performance, or cooperative production of responses by two or more speakers." Along similar lines, from her observation of "sharing time" classes, Michaels (1991) distinguishes between the ways that African American and European American children describe past events in their narratives. Further examining the same data as Michaels used, Gee (1991) points out the differences in the narrative techniques used by an African American girl and a European American girl; he categorizes the former as an oral-strategy (or poetic) narrative and the latter as a literate-strategy (or prosaic) narrative. These are some examples of cross-cultural comparison of narrative productions addressed in previous studies.

Furthermore, with regard to early language development at home and later language skills development such as literacy in school settings, the aforementioned illustrations predict different patterns of literacy development. In her ethnographic study, Heath (1983), for instance, describes children growing up in different communities – through interactions with adults around them – to be oriented toward particular genres and styles of narrative experiences. She suggests that "the ways of talking" information from text varies across cultures and social classes. When applied to the use of language, therefore, the social interaction paradigm suggests a culturally ideal adult-child relationship. Development takes place within the context of meaningful social interactions in which adults guide and scaffold children's participation in socioculturally appropriate ways. These examples illustrate the importance of considering cultural differences in the ways in which individuals structure their oral personal narratives, and these differences predict different literacy styles.

## 2 Maternal influence on the development of narrative discourse in Japanese children

Storytelling is an interactive social act that occurs in culture-specific patterns; within the aforementioned social interactionist paradigm (Bruner 1977; Vygotsky 1978), the research described in this chapter investigates the extent to which Japanese children's

personal narratives are influenced and constructed through social interactions. This paradigm maintains that talking about past events or experiences first takes place in interactive contexts, and that social support facilitates development in talking about the past (Sachs 1979, 1982).

Narrative is a superordinate term that includes many discourse genres, for example, personal anecdotes (i.e., autobiographical experience), fictional storytelling (i.e., pretend or role play), and scripts (i.e., the typical series of events that take place in a particular activity). Like other narrative genres, the telling of personal narratives is a universal element of human behavior, but such narratives embody culturally specific modes of telling a story. Personal narratives contain the elements with which individuals give accounts of their experiences in ways that they either consciously or unconsciously feel make sense within their culture. In other words, along with other possible factors, culture is part of what influences the elements that one selects or organizes for a narrative, such as causal logic, salience of a memory, and temporal sequence (McCabe and Bliss 2003).

The most basic requirement of a narrative is a recapitulation of chronologically sequenced events (Labov 1997, 2006) or, more generally, some reference to temporally or thematically connected events (Hicks 1994; McCabe 1991). Between the ages of two and three years, children's narrative structure moves from script-like accounts to specific recollections of real past events. English-speaking children begin to talk about the past at about two years of age (Eisenberg 1985). While children's language development progresses in this way toward extended narrative discourse, their early (through the age of three and a half years) productions are brief. Three-year-olds' narratives are often simple two-event narratives, whereas four-year-olds' narratives are much more diverse. Five-year-olds tell lengthy, well-sequenced stories that progress to the climax of the story more rapidly than adults (Dickinson and McCabe 1991; Peterson and McCabe 1983) regardless of what language they speak. In this way, young children's narrative skills continue to develop throughout toddlerhood and the preschool years in particular. They structure their oral personal narratives in an increasingly refined and mature style.

Previous research has raised a wide range of issues in the areas of narrative development. Hudson and Shapiro's (1991) work, for instance, treats how and what types of narrative develop at specific age levels of child development. Their study examines how the conversational context influences the coherence of children's narratives. Examining how the selection of the topic affects children's narrative production, Hudson and Shapiro illustrate how the components of narrative differ among children of different ages. Similarly, in her longitudinal study of three young children, Preece (1987) found that preschoolers are capable of producing a striking variety of narrative forms, such as personal anecdotes, parodies, film retellings, and fantasies. Over half of their conversational narratives, however, concern real personal experiences.

Research in narrative discourse has also examined the role of social interaction in young children's language development and acquisition of narrative structure. As stated earlier in the literature review section, it has been demonstrated that children's cognitive skills first develop through interactions with more mature members of society and are then internalized (Bruner 1977; Vygotsky 1978). In narrative contexts in particular, children's speech is guided and scaffolded by mothers who initiate topics and conversations. Hudson (1993), in her studies of the role of parent-child conversations in the development of young children's ability to talk about past events, investigated the effects of maternal elicitation styles. Her findings point out the influence of repetition in recounting events on the emergence and development of early autobiographical memory. Peterson and McCabe (1992) claim that stylistic differences between parents also affect children's later narrative style. Similar evidence comes from Fivush (1991) who suggests that those mothers who use a more elaborated elicitation style early in development raise children who provide more elaborated accounts later in their development. In fact, what Fivush and Fromhoff (1988) call "elaborative" mothers (those who provide a substantial amount of information) correspond to what Hudson (1993) calls "high elaboration mothers"; Fivush and Fromhoff's "repetitive" mothers, on the other hand, correspond to Hudson's "low elaboration mothers".

The aforementioned studies not only explored how maternal conversational styles support children's narrative development, but they also examined how maternal styles influence children's recounting past experiences in later years. While the way in which a mother's verbal interaction with her child may be reflected in the developing narrative skills of the child (Minami 2001), we must note that we cannot claim that parental narrative styles cause differences in narrative styles of their children. To some degree, we may be able to claim that the sorts of questions mothers ask during children's narratives predict the degree and characteristics of elaboration children incorporate into their narrative style. To be exact, however, we should note that there might exist discrepancies between the profile of parental narrative input and children's narrative productions.

Similar relationships on the role of mothers' linguistic responsiveness to young children have been reported in different cultures as well. For example, through her study of Mandarin-speaking children and their mothers living in Taipei, Taiwan, Chang (2003) claims that maternal approval to the child's talk, elaborative requests, and provision of information are correlated with children's narrative ability, such as describing objects and events and telling stories. Chang (2006) further claims that the existence of the continuous and interrelated relationship between early oral narrative and later language skills such as literacy is evident not only in English-speaking children [as reported, for example, by Snow and Dickinson (1991) and Tabors, Snow, and Dickinson (2001)] but also in Mandarin-speaking children. While any causal implications should be avoided, in narrative contexts children's speech is guided and scaffolded by mothers who initiate and elicit the children's recounting of

past experiences. In other words, parental talk conceivably provides a verbal framework for children's representations of past events. Because this view may imply that early social interactions shape young children's narratives into culturally preferred patterns, it is critical to consider cultural differences in how children structure their oral personal narratives.

Our aim here is to heighten awareness of the importance of narratives in the development of the first language (L1), Japanese in this case. It offers a unique approach to the study of narrative development that includes a significant cross-cultural dimension. The study specifically explores how language shapes and is shaped by culturally specific experiences through analyses of (I) how young children develop narrative structure, and (II) how parents guide their children in the acquisition of culturally appropriate styles of narrative and literacy. Studies have revealed that the succinct narrative style exhibited by Japanese children shows a remarkable contrast to the narrative style of North American children, which is typically a lengthy story detailing a single experience that often revolves around the solution of some problem (Minami 2002). Despite follow-up questions that encourage them to talk about one personal event at length, Japanese children generally present free-standing collections of several experiences; in contrast, North American children's narratives show a one-event-one-story scheme. To explore these differences, this study examines both children's monologic narratives and mother-child interactions or dialogic narratives.

The first topic of this L1 study relates to children's narrative discourse styles with particular attention to developmental features. The second topic of the L1 study explores the role of parental input in facilitating the development of children's personal narratives, based on the belief that the origins of narrative style can be traced back to early conversations at home between parents and children. These two topics provide a framework for consideration of some important issues regarding sociocultural contexts in L1 narrative discourse and even in foreign- or second-language (L2) narrative discourse that will be mentioned toward the end of this chapter.

## 2.1 Study one: Japanese children's narrative structure (monologic narrative)

### 2.1.1 Method

#### 2.1.1.1 Participants

The first study was designed to examine developmental patterns and culture-specific monologic narrative styles. Twenty middle-class preschool children in Japan participated in this study, along with their mothers. Of these children, 10 were in five-year-old children's classrooms in preschool (5 boys and 5 girls, $M$ = 5;03 years), whereas

the other 10 were in four-year-old children's classrooms (5 boys and 5 girls, $M$ = 4;03 years). Of the 20 children, 8 were first-born, 9 were second-born, and the remaining 3 were third-born. Among the 8 first-born children, 3 were only children at the time of data collection. All of the families were middle class as measured by the occupation and education of the father, and the education of the mother.[4] Of the 20 fathers, 16 had college degrees, 2 had graduated from professional school. One father had finished high school. One father had attended graduate school. Of the 20 mothers, 12 had attended two-year (junior) college. One had attended professional school, 2 had concluded their schooling upon graduating from high school, and 5 had four-year college degrees.

In separate interviews, these children and their mothers were prompted with questions related to injuries or early childhood memories, in the manner employed by Peterson and McCabe (1983). All subjects were monolingual, and none of these mother-child pairs had experienced living overseas at the time of data collection.

#### 2.1.1.2 Task, procedure, and materials

This activity, which analyzes children's developing narrative skills, follows the methodology for eliciting personal narratives developed by Peterson and McCabe (1983) for use with young children. Before eliciting narratives, rapport was established with the child through activities such as drawing pictures.[5] When children were judged to be comfortable in making conversation with the interviewer, they were asked prompting questions related to personal experiences (injuries for children, and early childhood memories for mothers). This elicitation technique had previously proved to be effective with Japanese children (Minami 1996a, 1996b; Minami and McCabe 1995). Questions were asked about personally experienced events, such as whether they had ever gotten hurt.[6] During the interviews, only nonspecific social support was provided; that is, the interviewer's responses were confined to repeating the exact words of the children during a pause, back-channels (1a), and nonspecific requests and questions (1b) and (1c).

**4** Japanese parents of different socioeconomic status might guide their children in the acquisition of socioculturally specific styles of narrative.

**5** To minimize the child's self-consciousness and not to influence his or her social behaviors, the conversations were recorded at the child's home or, if it was not available for some reason, the child's friend's home was used as an alternative. In this sense, interactive interviewing – "a conversational, open, or loosely structured mode, in which give and take is emphasized" (Modell and Brodsky 1994: 142) – was conducted.

**6** Injury is a topic that typically elicits extensive narrative production even from very introverted or young children (McCabe and Peterson 1991; Peterson and McCabe 1983).

(1) a. *un, un.* *huun.*
'uh huh.' 'well.'

b. *motto hanasite kureru?*
'Would you tell me more?'

c. *sorede doo natta no?*
'Then what happened?'

Note that individual differences in elaborative style may have been positively correlated with the rate of prompts during the story, and if this were the case, one might argue that the more elaborative storytellers are, the less dependent they are on prompts. However, although there are differences in general subprompts (e.g., the back-channel "uh huh" and more specific questions), all subprompts are statements or questions that do not refer to the content of a story but serve only to encourage the narrator to continue talking. In other words, these responses are relatively neutral and simply serve as signals of the interviewer's interest to the children. Furthermore, parents did not accompany their children while the interviews were in progress. The narratives elicited from the children were thus minimally scaffolded and relatively *monologic* in nature.

As with the children, an attempt was made to elicit monologic narratives from adult speakers, mothers in this study. Unlike the children, however, because the adults were expected to express themselves more easily, such tasks as drawing pictures were not used as warm-up tools. Unlike children, furthermore, because adults were expected to express themselves with little difficulty, the injury story was not used as an elicitation technique. Instead, they were asked to talk about any experiences, such as their earliest memory.

To understand language-specific styles of narrative discourse, we need to consider the differences in how young children and more mature speakers of the same language tell narratives. To examine the relative narrative competence of young children, therefore, narratives were also elicited from adults – the children's mothers in this study. In the past, only a few narrative researchers have had adults perform the same task as children (e.g., Berman and Slobin 1994). Few studies of Japanese have extensively examined adults' narrative style, although Maynard (1993), for instance, identified a variety of Japanese linguistic devices and manipulative strategies (e.g., evidentiality morphemes [particular verb-ending forms/suffixes], modal adverbs, and discourse connectives) that convey a subjective emotion as well as an individual's shared feelings with others.[7] Analyzing adults' narratives, however, is critical in that

7 Maynard (1993) analyzes conversational data as well as narrative segments, which she took from contemporary works of Japanese fiction. She explores multiple functions of narrative discourse connectives *dakara* (so, therefore) and *datte* (but, because) in interactional contexts. For instance, the narrative segments subsequent to *dakara* provide the listener with supplementary information; the narrator conveys his or her personal information, attitude, and emotion toward the fact in

we can examine what linguistic devices and narrative strategies preschoolers have not yet begun to deploy in their narratives as compared to adults. Thus, adults' narratives are assumed to serve as a standard of comparison that provides the culturally appropriate, full-fledged, rhetorically well-formed narrative that children are expected to accomplish eventually in telling personal narratives.

#### 2.1.1.3 Coding

One piece of research relevant to the present study is the seminal study conducted by William Labov (1972), the sociolinguist who pioneered the study of oral personal narratives through the examination of the interface between cultural and linguistic issues. According to Labov, the narrator relies on affective expression as a primary means of conveying the relational significance of narrative events. A coding scheme has been developed using Labovian methodology (Labov 1997) to interpret the data focusing on the content of each clause in monologic narrative from the standpoint of *high point analysis* (i.e., the narrative event is considered to culminate in a high point of some kind).

Although different scholars had proposed other analytic units in the past, Labov and Waletzky (1967) adopted the independent clause for their analyses, and designed linguistic techniques to evaluate the narration of experiences in African American Vernacular English. According to Labov (1972), a narrative consists of two important elements: referential and affective (or evaluative). The referential elements, which convey information about events and characters, are further categorized into two components: complicating action and orientation. Complicating actions depict the sequence of specific, chronologically ordered events comprising the experience, whereas orientations, unlike specific actions, provide descriptive non-sequential information, setting the stage for the narrated events, such as information about people, place(s), time(s), and situation(s). In other words, complicating actions give plot-advancing foreground information, whereas orientations provide contextualizing background information of the story (Hopper 1979; Hopper and Thompson

---

a somewhat brief way. Giving detailed descriptions of the roles of modal adverbs, Maynard (1993: 127) also emphasizes that an adverb *yahari/yappari* (anyway, after all) "brings with it a feeling of speakerhood"; that is, *yahari/yappari* "attests to the fact that the utterance is personalized and the speaker is there." Maynard further examines differences in narrative function between the plain form of the copula *da* (a linking verb "be") and formal *des(u)/mas(u)* verb ending forms (i.e., suffixes) in Japanese [note that *des(u)* is the polite form of abrupt *da*]. Although Maynard does not elaborate on the relation between narrative point of view and interactional particles, she does explore the functions of two interactional particles: (a) assertive and/or emphatic particle *yo*, which, being similar to English expressions "I tell you" and "I'm sure", expresses the speaker's insistence or forcing the given information on the addressee; and (b) a rapport particle *ne*, by which the narrator seeks the listener's agreement as English speakers use "you know", "right?", "don't you agree", or tag question [in other words, *ne* (or *nee* when elongated) expresses a request for compliance with the given information leaving the option of confirmation to the addressee].

1980). Evaluative elements, which also provide background information, convey the narrator's attitudes toward events and his or her interpretations of the protagonists' motives and reactions to events. Without evaluations, narrated events would be mere representations of facts of our lives.

In addition, two other categories were also included: (i) appendages, such as abstracts ("What, in a nutshell, is this narrative about?") and codas ("That's it", which signals the *sealing off* of a narrative), and (ii) reported speech (statements that report character speech by generally reproducing the speech performed). Note that reported speech, which is considered a linguistically marked recounting of a past speech event, is an important category because second-order evaluations (which will be explained soon) are possibly provided in reported speech (Ely and McCabe 1993; Gwyn 2000).

Minami and McCabe (1995) used an earlier adaptation of Labov's technique referred to as high point analysis by Peterson and McCabe (1983) because of the critical importance of ascertaining the emotional climax – the high point – for evaluating the narrative as a whole. The coding scheme specifies the role that each clause plays in the organization of a narrative. In other words, personal narratives are fundamentally accounts of past experiences, but simply narrating what happened is not enough. Instead, in order to establish an appropriate spatial and temporal context, different types of information need to be encoded in a narrative. For example, some clauses play the role of descriptive "orientation", which is considered to be the stage setting for the narrated events, whereas others function as "evaluation" or represent "actions" (or "complicating action" to follow Labov's terminology). To summarize, most clauses produced by the participants were coded into one of the following categories:

(2) *Complicating action*: Specific actions, events, or processes that take place, which are thus temporally *restricted* narrative clauses. An "action clause" recapitulates a single event that took place at some discrete or restricted point in time.
e.g., *tyuusya sita.*
'[I] got an injection.'

(3) *Orientation*: Statements in which the narrator digresses from the events of a narrative to provide the listener with contextual embedding, such as features of environment, conditions, and ongoing behavior in the narrative. Orientation clauses do not occur at a restricted point in time and are thus relatively *free* narrative clauses.
e.g., *tyuurippu gumi no toki.*
'When [I] was in the Tulip class at my preschool.'

(4) *Evaluation*: Another type of relatively *free* narrative clause that tells the listener what to think about a person, place, thing, event, or more globally, the entire experience described in a narrative telling.
e.g., *nakanakatta.*
'[I] didn't cry.'

(5) *Appendage*: A composite category including narrative comments that appear either at the beginning (abstracts, attention-getting devices) or at the end of the main body of the narrative (codas).
Abstracts: e.g., *kaze no toki mo itta koto aru.*
'Even when [I had] cold, [I] went.'
Attention-getting devices: e.g., *boku sitteru wa.*
'I know [something].'
Codas: e.g., *osimai.*
'That's it.'

(6) *Reported speech*: Statements that report character speech by generally reproducing the speech performed.
e.g., *ano ne, "tyuusya sinai hito ehon yonde mattoite kudasai," tte* sensei itteta.
'Well, you know, "Those who haven't gotten a shot, please read a book while you're waiting," said the teacher.'

The coding scheme was piloted on a portion of the data. After the principal coder had analyzed all 40 transcripts of the monologic narrative data, a second coder independently analyzed 8 of those transcripts – 4 of the 20 adult transcripts and 4 (2 from each age group) of the children's transcripts. Inter-rater reliability as measured by Cohen's kappa, which is an estimate of reliability that corrects for chance rates of agreement, was .94 for the main categories (i.e., complicating action, orientation, evaluation, appendage, and reported speech) of the children's coding, and .95 for the main categories of the adults' coding, representing "almost perfect" agreement (Bakeman and Gottman 1997; Landis and Koch 1977). Once all the transcripts were coded, a series of *C*omputerized *L*anguage *An*alysis (CLAN) programs (MacWhinney 2000) was employed to analyze frequencies of different codes, the total number of words, the number of different words, and the total number of clauses.

### 2.1.2 Results

Using the aforementioned coding rules, several principal characteristics of Japanese personal narratives were quantitatively analyzed.

#### 2.1.2.1 Narrative length

Narrative length was measured in two different ways: the total number of words in the narrative and the total number of subject-predicate clauses. For the four-year-olds, the total number of words ranged from 31 to 129 ($M$ = 75.90, $SD$ = 31.77), the total number of different words from 19 to 65 ($M$ = 38.10, $SD$ = 13.76), and the total number of clauses (i.e., subject-predicate propositions) from 5 to 16 ($M$ = 10.90, $SD$ = 3.60). For the five-year-olds, the total number of words ranged from 32 to 161 ($M$ = 75.60, $SD$ = 47.80), the total number of different words from 18 to 64 ($M$ = 37.10, $SD$ = 17.76), and the total number of clauses from 7 to 19 ($M$ = 10.70, $SD$ = 3.95). Additionally, similar type-token ratios (the total number of different words divided by the total number of words: Templin 1957) of these two age groups – .52 for four-year-olds and .53 for five-year-olds – indicate that the levels of lexical redundancy are almost identical. No age differences approached statistical significance in the three categories: (a) the total number of words, $t(18) = .02$, *ns*, (b) the total number of different words, $t(18) = .14$, *ns*, (c) the total number of clauses, $t(18) = .12$, *ns*, and (d) the type-token ratio, $t(18) = .31$, *ns*. As far as these production variables are concerned, therefore, no major developmental differences between four-year-olds and five-year olds were observed (see Table 1 below).

**Table 1:** Means and Standard Deviations of Total Number of Words, Total Number of Different Words, Type-Token Ratios, and Total Number of Clauses in Monologic Narrative Production

| | Mothers of four-year-olds ($n$ = 10) | | Mothers of five-year-olds ($n$ = 10) | | | |
|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *t* value | *df* |
| Total number of words | 75.90 | 31.772 | 75.60 | 47.796 | .017 | 18 |
| Total number of different words | 38.10 | 13.763 | 37.10 | 17.760 | .141 | 18 |
| Type-token ratio | .521 | .088 | .533 | .092 | .309 | 18 |
| Total number of clauses | 10.90 | 3.604 | 10.70 | 3.945 | .118 | 18 |

What attracts our attention, however, is that, in these three categories (i.e., the total number of words, the total number of different words, and the total number of clauses), no significant associations were found between mothers and children. Pearson's product-moment correlation between mothers and children for the total number of words used was $r(18) = .19$, *ns*. The correlation for the total number of different words used was $r(18) = .22$, *ns*. Likewise, the correlation for the total number of different words used was $r(18) = .30$, *ns*. These results thus indicate that talkative mothers do not necessarily have talkative children, and that reticent mothers do not necessarily have reticent children.

#### 2.1.2.2 High point analysis

As a content-based narrative analysis, narrative elements (i.e., complicating action, orientation, evaluation, appendage, and reported speech) were also examined. No differences reached statistical significance between four-year-olds and five-year-olds, in terms of complicating action, evaluation, or orientation in either raw or proportional frequencies.

Each of these two age groups was then compared with the adult group (i.e., 20 mothers). The proportion of children's *restricted* narrative clauses (i.e., complicating actions) was substantially higher than adults'. That is, four-year-olds provided proportionately more action statements than adults (i.e., mothers), $t(28) = 5.21$, $p < .0001$; five-year-olds also provided proportionately more action statements than adults, $t(12) = 3.58$, $p < .004$.[8]

Conversely, the proportion of the sum of evaluation and orientation clauses (i.e., background information) provided by adults was substantially higher than that by either one of the children's age groups. That is, the proportion of the narrative that four-year-olds devoted to background information was lower than the proportion of the narrative that adults (mothers in this case) devoted to background information, $t(28) = -4.53$, $p < .0001$. Also, the proportion of the narrative that five-year-olds devoted to background information was lower than the proportion of the narrative that adults devoted to background information, $t(28) = -3.43$, $p = .002$.

The two types of background information (i.e., orientation and evaluation) were then examined separately because orientation serves a referential function but evaluation serves an affective or evaluative function (Labov 1972). A difference was observed between four-year-olds and adults in orientation statements, $t(28) = -2.79$, $p = .009$. Likewise, the proportion of the narrative that five-year-olds devoted to orientation statements was lower than the proportion of the narrative that adults (mothers in this case) devoted to orientation statements, $t(28) = -3.22$, $p = .003$. An age related difference, however, was identified. The proportion of the narrative that four-year-olds devoted to evaluation statements was lower than the proportion of the narrative that adults devoted to evaluation statements, $t(28) = -3.02$, $p = .005$, but the difference between five-year-olds and adults did not reach a statistical significance in terms of proportional frequencies, $t(12) = -1.00$, *ns*. The evaluation revealed that five-year-olds' monologic narrative organization, compared to four-year-olds' one, tends to resemble adult models.

**8** Note that because the length of adults' narratives (i.e., the total number of clauses) are significantly longer than that of four-year-olds, $t(28) = 4.93$, $p < .0001$, and that of five-year-olds, $t(28) = 4.95$, $p < .0001$), comparing raw frequencies, particularly in terms of examining narrative structure, is meaningless. Note also that whenever the population variances are not assumed to be equal, the *t*-statistic based on unequal variances was used. When the variances of the two samples were quite different, therefore, this procedure reduced the degrees of freedom.

To summarize, Labovian methodology (Labov 1997) is associated with content-based orientations to narrative. Labov and Waletzky (1967) elicited narratives of a life-threatening experience from adolescents in Central Harlem, New York City. Two important features characterize Labovian methodology. First, its focus is on the temporal sequencing of the linguistic string as critical to narrative accounts of events. This is evident in Labov and Waletzky's (1967: 20) earlier definition of narratives – "one method of recapitulating past experience by matching a verbal sequence of clauses to the sequence of events which actually occurred",[9] as well as Labov's much later explanation:

(7) The fundamental concept that distinguishes narrative from other ways of reporting the past is temporal juncture: a relation of before-and-after that holds between two independent clauses, and matches the order of events in time. Such sequences of ordered clauses form the complicating action that is the skeletal structure of narrative.
(Labov 2006: 37)

The second important feature is the distinction between two elements required for a successful narrative: referential elements and evaluative elements.

The analysis of the responses was done using Labovian methodology (Labov 1997), and the results indicated the following: (a) Children tend to tell their stories in a sequential style (i.e., action statements: foreground information), whereas adults emphasize non-sequential, background information (i.e., evaluation or orientation statements). (b) Both four-year-olds and five-year-olds emphasize a simple description of successive events (i.e., foreground information). Compared to four-year-olds, however, five-year-olds begin to provide evaluative comments (i.e., background information) in ways that are slightly more adult. Narrative competence derives from a cognitive schema that is shared across mature speakers. It requires knowledge of core plot components, or what Labov (1972) termed referential elements and devices for evaluation so that narrators alternate plot-advancing foreground information and contextualizing background information for sophisticated elaboration. Overall, the results provide strong evidence that the following clear differences exist in terms of content and delivery between children and adults: (a) Children tend to tell their stories in a sequential style, whereas adults emphasize non-sequential information. (b) With age, however, children steadily and increasingly include non-sequential information in order to achieve greater narrative coherence, which is the central aim of most adult narrators.

**9** See also Labov (1972: 359–360).

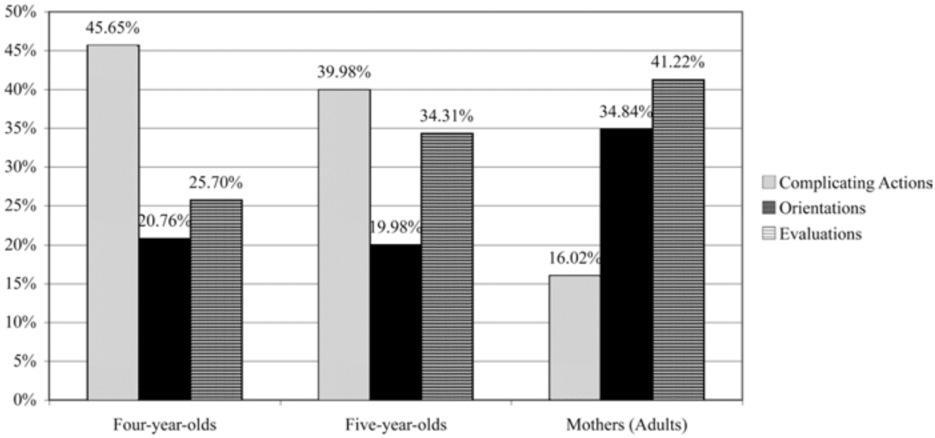



**Figure 1:** Distribution of the three major narrative components: A comparison of preschoolers and their mother's monologic narratives

## 2.2 Study two: Japanese mother-child joint narratives (dialogic narrative)

### 2.2.1 Age differences within a particular culture

Content-based narrative analyses offer important predictions for the development of narrative abilities. Labovian methodology (Labov 1972, 1997), which belongs to content-based narrative analyses, is considered appropriate for the current study, because it focuses on "such factors as the event representation underlying the prototypical situation or script in which a narrative is anchored, the narrative as a text leading up to a high point, and embedding the sequential chain of events in a network of evaluative comments and background circumstances" (Berman 1995: 287–288). To begin with, as Nelson (1986, 1989) suggests, for successful narrative production young children need to be cognizant of a familiar script (i.e., verbal reconstruction of familiar sequences serves as the foundation of verbal reconstruction of personal experience). Furthermore, as Peterson and McCabe (1983) emphasize, child narrators are required to have a good command of canonic narrative structure (i.e., a series of complicating actions leading up to a highpoint and culminating in a final outcome or resolution). Finally and most important, child narrators – young children in particular – may provide referential narrative information (e.g., plotline events) but they may not provide the amount of evaluative interpretation that is necessary for successful storytelling. In other words, young children tend to focus on events and activities but they may not devote an ample amount of verbal expression to motivational, evaluative, and other background information (Berman and Slobin 1994).

In this context, we need to consider the role of scaffolding. The ideal goal is telling narratives virtually without any support (i.e., monologic narrative). In reality, however, young children develop narrative abilities in essentially interactive contexts. That is, when storytelling is clearly interactive and embedded in conversation (i.e., dialogic narrative), with supportive input from familiar adults, mothers in many cases, children are likely to produce well-constructed strings of narrative discourse. Responding to scaffolding input, young children construct a story as part of dialogic interchange. Because of this, while the first question addressed young children's freestanding personal narratives, the second question looked at their narratives in the context of mother-child interactions. This decision was thus made because narratives are, after all, dialogically evolving episodes of interaction, in which evaluations are frequently co-constructed between speaker and listener (Gwyn 2000). The same 20 Japanese mother-child pairs were used to investigate this question. However, unlike the interviews described above, in this data collection mothers were asked to tape-record, at home, conversations with their children about past experiences.

### 2.2.2 Method

#### 2.2.2.1 Task, procedure, and materials

As reviewed earlier, previous research (e.g., Fivush 1991; Hudson 1993) reported that some mothers adopt an elaborative (or high elaborative) style, whereas other mothers adopt a repetitive (or low elaborative) style engaging in conversations during which they provide little descriptive information. The purpose of examining parental styles of narrative elicitation is, in accordance with Vygotskian sociocognitive theoretical models (Vygotsky 1978), to understand mothers' scaffolding strategies, i.e., how mothers verbally interact with their young children during narrative elicitation. A great number of narratives about children's experiences are, in fact, told jointly by parents and children; that is, the parent prompts particular information from the child as well as adds information (Peterson 2004). A considerable number of research (e.g., McCabe and Peterson 1991) has revealed that such co-construction teaches children how they should structure their narratives and what kinds of information they should include.

In this study, mothers were asked to elicit talk about interesting past events or experiences from their children in a relaxed and informal situation; no other specific instructions or requests were provided. In contrast to the researcher's elicitation strategies in which only non-directive general cues were given and thus narratives were elicited from the children with minimal scaffolding (as described in the "children's narrative structure" subsection above), in this activity mothers were expected to scaffold the narratives of their young children. Here, following the methods developed by McCabe and Peterson (1991), mothers were encouraged to

ask their children to relate stories about personal experiences, about real events that had happened in the past. However, mothers were also asked to do this narrative elicitation in as natural a way as possible, just as they would ordinarily behave when asking their children to talk about past events. In contrast to the researcher's narrative elicitation, this task is thus *dialogic* in nature. The dialogues between mothers and their children were recorded in their homes.

#### 2.2.2.2 Coding

Transcripts of all parents' speech were scored according to a coding system that was previously used to analyze how speech acts are mapped onto dialogic narrative discourse in English (McCabe and Peterson 1991). Using this coding scheme as a basis, Minami and McCabe (1995) devised coding rules that are applicable to Japanese data. All parental speech was scored according to these coding rules. Parental utterances were coded as one of three types: (I) topic-initiation (or topic-switch), (II) topic-extension, and (III) conversational strategies that simply show attention, such as "*un*" ('uh huh') and "*huun*" ('well'). Utterances categorized as topic-extension are further subdivided into:

(8) *Descriptive statements*: Statements that describe or request description of a scene, a condition, or a state.
e.g., *ato Momotaro no hon mo atta desyo.*
'There was also a book about the Peach Boy.'

(9) *Action statements*: Statements about or requests for information about complicating actions that, accompanied by an action verb, describe a specific action.
e.g., *banana mo tabetan.*
'[You] also ate a banana.'

(10) *Mother's evaluative comments*: Evaluations by the mother herself.
e.g., *sore ii ne.*
'That's good, wouldn't you say.'

(11) Mother's request for child's evaluative comments:
e.g., "*uu-tyan no doko ga kawaii no?*"
'What do [you] think is cute about the bunny?'

The coding categories used are (i) maternal requests for the child's descriptions, complicating actions, and evaluations, (ii) maternal evaluations, and (iii) statements showing attention. Note that in this study, because the number of initiation was controlled, it was not included.

This study, analyzing the connection between maternal narrative elicitation strategies and children's developing narrative skills (which were described in the "children's narrative structure" section), specifically focuses on the following correspondence between monologic and dialogic narrative: (a) the relationship between the child's action statements in monologic narrative and the mother's statements about or requests for information about complicating actions during narrative elicitation; (b) the relationship between the child's orientation statements in monologic narrative and the mother's descriptive statements that describe or request description of a scene, a condition, or a state; (c) the relationship between the child's evaluation statements in monologic narrative and evaluation components during maternal narrative elicitation, i.e., (i) the mother's evaluative comments and (ii) the mother's request for the child's evaluative comments. Emphasis on these two-way interactions indicates the belief that children and their environments – mothers in this case – need to be conceptualized as a dynamic system in which they actively interact with and influence each other. This bidirectional emphasis furthermore plays a complementary role for the Labovian analysis (Labov 1972, 1997) in which, at the expense of interactional aspects, the structural dynamics of storytelling are emphasized (Gwyn 2000). In evaluation in particular, the integration of bidirectional interactions lets the Labovian account of evaluation explain the interactional functions of narrative (Wortham 2000).

### 2.2.3 Results

The raw frequency of each category of parental speech was calculated. The proportional frequency was also determined by dividing the total frequency of each category by the total number of utterances that the mother produced. Frequencies were analyzed because they represent the impact that loquaciousness might have on children's narration (e.g., Hoff-Ginsberg 1992). Note that proportions were also used because they correct for differences in length and allow us to see differing relative emphasis on components of narration. But no statistically significant differences were observed from proportional frequencies in this case.

To evaluate whether mothers of four-year-old children and mothers of five-year-old children elicited in different ways, a series of independent-samples *t* tests was conducted for the major coding categories: maternal requests for descriptions, complicating actions, and evaluations, maternal evaluations, and statements showing attention. The following differences due to the age of the children emerged in mothers' narrative elicitation strategies: (a) Mothers of four-year-olds requested evaluation from their children more frequently than did mothers of five-year-olds, $t(18) = 3.18$, $p = .005$. (b) While *t* tests for the other coding categories did not reach statistical significance, compared to the mothers of five-year-olds, the mothers of four-year-olds were more likely to give their children topic-extension prompts (see

Table 2). It should be noted that when adult interviewers elicited narratives from the children without providing such scaffolding (see the "children's narrative structure" subsection above), four-year-olds provided less evaluation than adults, while no differences were observed between five-year-olds and adults. The observed difference in maternal narrative elicitation patterns therefore shows that the mother guides her child by providing requests for evaluation if she considers it appropriate unless mothers judge that evaluation statements are sufficient.

**Table 2:** Mean Frequencies and Percentages (Standard Deviations) of Mothers' Prompts to Children about Past Events

| | Mothers of four-year-olds ($n$ = 10) | | Mothers of five-year-olds ($n$ = 10) | | | |
|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *t* value | *df* |
| Requests for descriptions | | | | | | |
| Frequencies | 19.10 | 8.937 | 15.00 | 7.348 | 1.121 | 18 |
| Percentages | 13.71% | 3.533 | 15.18% | 5.025 | .756 | 18 |
| Requests for actions | | | | | | |
| Frequencies | 28.80 | 15.591 | 23.50 | 10.014 | .905 | 18 |
| Percentages | 20.75% | 9.083 | 25.09% | 11.762 | .924 | 18 |
| Requests for evaluations | | | | | | |
| Frequencies | 32.00 | 13.241 | 16.50 | 7.878 | 3.181** | 18 |
| Percentages | 25.40% | 11.190 | 17.73% | 8.619 | 1.716 | 18 |
| Evaluations by mother herself | | | | | | |
| Frequencies | 22.70 | 14.606 | 15.40 | 10.341 | 1.290 | 18 |
| Percentages | 15.33% | 6.208 | 15.10% | 6.721 | .079 | 18 |
| Statements showing attention | | | | | | |
| Frequencies | 36.20 | 31.460 | 27.10 | 20.851 | .762 | 18 |
| Percentages | 24.81% | 17.804 | 26.89% | 19.607 | .248 | 18 |

**$p$ < .01.

## 2.3 Study three: Cross-cultural comparison of parental styles of narrative elicitation

Mother-child storytelling has been used as a first step toward exploring language socialization through narrative discourse. Because storytelling is an interactive social act that occurs in culture-specific patterns, it is not surprising that a large amount of cross-cultural research has been conducted in the context of parent-child narratives, in part with the intent of revealing the manner in which culture is reflected in narratives (Harkins and Ray 2004; Shulova-Piryatinsky and Harkins 2009). However, the studies that have investigated sociocultural variations in maternal narration in style are still few (Melzi and Caspe 2005).

The purpose of this study is to address this gap through an exploration of how mothers from different cultural groups engage in joint narratives. As a continuation of Study Two, the research was then extended using a cross-linguistic/cross-cultural approach to compare narrative elicitation patterns in different language/cultural groups. The study compared conversations between mothers and children from three different groups: (a) Japanese-speaking mother-child pairs in Japan (the mother-child pairs described in the previous sections), (b) Japanese-speaking mother-child pairs living in the United States, and (c) English-speaking North American mother-child pairs. Note that the research presented here deals exclusively with the comparisons of five-year-olds: (a) 10 middle-class Japanese five-year-olds (5 boys and 5 girls, $M$ = 5;03 years) and their mothers living in Japan (none of these mother-child pairs had experienced living overseas at the time of data collection), (b) 8 middle-class Japanese five-year-olds (4 boys and 4 girls, $M$ = 5;03 years) and their mothers living in the United States, and (c) 8 English-speaking middle-class North American five-year-olds (4 boys and 4 girls, $M$ = 5;03 years) and their mothers. All mothers agreed to interview their children at home about real past events.

### 2.3.1 Method: Task, procedure, materials, and coding

As previously described in the "mother-child interactions and children's narratives" subsection, mothers were asked to elicit narrations from their children about interesting past events or experiences in a setting that was in every way relaxed and informal; no further specific instructions or requests were provided. Subsequently, the speech act coding system described in a previous section was also used for analyzing the parental styles of narrative elicitation. Speech was broken into utterances, and transcripts of all parents' speech were scored according to that system.

### 2.3.2 Results

#### 2.3.2.1 Maternal styles of narrative elicitation

For parental speech, the total frequency of each category was counted, and also the proportional frequency was calculated by dividing the total frequency of each category by the total number of utterances that the mother produced. To test for the effect of *group* (Japanese mother-child pairs in Japan, Japanese mother-child pairs in the United States, and North American mother-child pairs), multivariate analyses of variance (MANOVA) were conducted for the major coding categories: maternal requests for the child's descriptions, actions, and evaluations and maternal evaluations and statements showing attention (see Table 3).

**Table 3:** Mean Frequencies and Percentages of Mothers' Prompts to Children about Past Events (five-year-olds)

| | Japanese mothers in Japan ($n$ = 10) | | Japanese mothers in the U.S. ($n$ = 8) | | North American mothers ($n$ = 8) | | |
|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | $F^a$ value |
| Requests for descriptions | | | | | | | |
| Frequencies | 15.00 | 7.348 | 14.00 | 7.892 | 17.63 | 14.995 | .261 |
| Percentages | 15.18% | 5.025 | 21.91% | 4.704 | 19.54% | 5.454 | 4.127* |
| Requests for actions | | | | | | | |
| Frequencies | 23.50 | 10.014 | 15.50 | 10.915 | 17.88 | 15.932 | 1.009 |
| Percentages | 25.09% | 11.762 | 23.89% | 6.940 | 20.39% | 12.687 | .436 |
| Requests for evaluations | | | | | | | |
| Frequencies | 16.50 | 7.878 | 8.75 | 5.120 | 21.38 | 19.885 | 2.131 |
| Percentages | 17.73% | 8.620 | 15.07% | 9.639 | 22.20% | 9.803 | 1.206 |
| Evaluations by mother herself | | | | | | | |
| Frequencies | 15.40 | 10.341 | 7.75 | 10.846 | 28.25 | 25.822 | 3.070† |
| Percentages | 15.10% | 6.721 | 9.21% | 7.496 | 29.13% | 13.024 | 9.778*** |
| Statements showing attention | | | | | | | |
| Frequencies | 27.10 | 20.851 | 17.50 | 6.803 | 7.50 | 7.270 | 4.273* |
| Percentages | 26.89% | 19.607 | 29.92% | 9.840 | 8.74% | 4.204 | 5.795** |

†$p < .07$. *$p < .05$. **$p < .01$. ***$p < .001$.
[a] Degrees of freedom = 2, 23.

In terms of proportions, there was a significant multivariate effect of *group*, Wilks' lambda = .33, approximate $F(8, 40) = 3.69$, $p = .003$. Univariate ANOVAs were run for each of the dependent variables. The effect of *group* was largely attributable to significant effects on maternal evaluations, $F(2, 23) = 9.78$, $p = .001$, and statements showing attention, $F(2, 23) = 5.80$, $p = .009$. The results were further analyzed in Fisher's least significant difference (LSD) tests, which revealed the following: (a) North American mothers gave proportionately more evaluation ($M$ = 29.13%, $SD$ = 13.02) than did both Japanese mothers living in Japan ($M$ = 15.10%, $SD$ = 6.72) and in the United States ($M$ = 9.21%, $SD$ = 7.50). (b) On the other hand, both Japanese mothers living in Japan ($M$ = 26.89%, $SD$ = 19.61) and in the United States ($M$ = 29.93%, $SD$ = 9.84) gave proportionately more verbal acknowledgment (i.e., statements showing attention) than did North American mothers ($M$ = 8.74%, $SD$ = 4.20).

#### 2.3.2.2 Child's length of turns

In addition to the frequencies of the coded behaviors, the "child's utterances over turns" was examined ("utterances over turns" can be defined as the number of

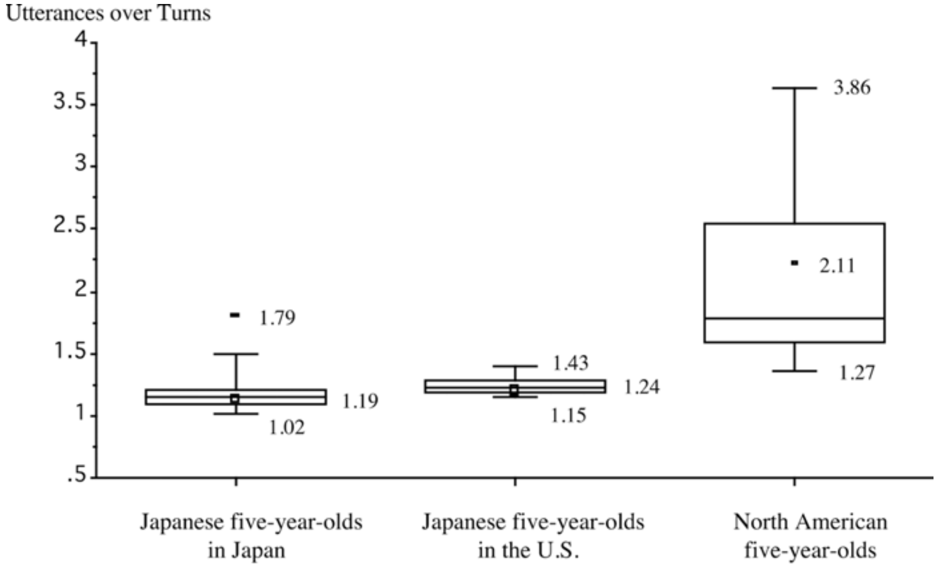



**Figure 2:** Children's ratio of utterances over turns

utterances produced by a speaker per turn).[10] North American children produced approximately 2.11 utterances per turn on average (*SD* = .90). On the other hand, Japanese children living in Japan and the United States produced 1.19 utterances (*SD* = .22) and 1.24 utterances (*SD* = .10), respectively. A one-way analysis of variance (ANOVA) was performed on the variable, "utterances over turns". This ANOVA yielded a significant main effect of *group*, $F(2, 23) = 8.34$, $p = .002$. The ANOVA results were further analyzed in the LSD Post Hoc tests, which showed that Japanese children, whether living in Japan or the United States, produced fewer utterances (i.e., about 1.22 utterances on average) per turn than did North American children (Figure 2). These differences in utterances over turns indicate differences in the direction of maternal control, i.e., to what extent mothers allow their young children to take long monologic turns.

To summarize, the results indicate: (a) English-speaking North American mothers, while allowing their young children to take long monologic turns, offer positive evaluation of the narrative (e.g., "It's a good story."); (b) Japanese mothers, on the other hand, indicate their interest in their children's narratives by frequently providing brief verbal signs of encouragement (e.g., "*un*" ['yes'] "*huun*" ['well']). They facilitate frequent turn exchanges while offering few evaluative comments. From

**10** Note that even if an *aizuchi* (which is similar to back-channeling) appears, that does not necessarily mark the end of a turn, because it may indicate an unconditional signal to go on talking.

early childhood on, children in all groups become accustomed to culturally valued narrative discourse skills through interactions with their mothers.

### 2.4 Summary and discussion for L1 narrative development

In this study, the genre of narrative (or story) was investigated in the larger context of discourse, from both developmental and cross-cultural perspectives. The ability to narrate a story is considered a skill, and the study first examined how Japanese children acquire and develop narrative discourse skills. According to Labov (1972), referential and affective (or evaluative) elements are both indispensable for an ideal narrative. Action statements serve to move the plotline forward as it proceeds from orientation statements, whereas affective (or evaluative) statements convey the narrator's attitudes toward the narrative events. Furthermore, sequential "foreground" information (i.e., action statements) refers to the parts of the narrative that relate a sequence of events with respect to a timeline and thus constitute the *skeletal structure* of the narrative. In contrast, non-sequential "background" information refers to supportive narrative (e.g., orientation to where and when events occurred, and evaluation, which describes the agent's motives) that does not itself relate the main events. The results reported in this research indicate that compared to adults, preschool children focus on temporal sequence with little emphasis on non-sequential information such as evaluation and orientation.

In an effort to comprehend children's acquisition of a culturally shared narrative style, this study also examined how the culture-specific aspects of young children's narrative skills are developed through mother-child conversational interactions. The study specifically examined cultural variations in parental goals of storytelling and story constructions to and with young children. It has revealed that mothers in each culture simultaneously pay considerable attention to all the elements in their children's narratives and, through specific types of intervention (i.e., North American mothers allow their children to take long monologic turns, and subsequently give many evaluative comments, whereas Japanese mothers intervene with frequent turn exchanges), actively support the progressive development of their children's narrative skills. As summarized in the previous two subsections, maternal styles of narrative elicitation may affect their children's narrative techniques later.

## 3 From present research to future research

### 3.1 Overall summary

The progression of language growth is remarkable as children move from expressing their needs and emotions to adults through cries and babbling sounds, to the rapid expansion of language, as children and adults (mainly their parents), through verbal means, engage in on-going activities, and eventually to children and adults sharing

experiences as they relay personal narratives. “How do children learn to narrate?” is one of the questions posed in this chapter. Sometime during the second year of life – anywhere from 12 to 18 months – children begin to utter their first words. During the following four to five years, language acquisition and development occur quite rapidly. For example, when Labovian methodology (Labov 1972, 1997; Minami 2002) is applied to the narrative of Akio, a boy aged 5 years and 4 months, the following organization is identified. To present his injury experience in this narrative, Akio mostly used chronological statements in a time sequence. In other words, while he chained events sequentially and structured his narrative causally, he did not necessarily elaborate on background information such as emotional states.

(12) Akio’s monologic narrative

*Orientation*

*ano ne, supiido dasite itara ne,*

‘Um, you know, (I) was speeding, you know,’

*Complicating action*

*koronde,*

‘(I) fell,’

*Complicating action*

*handoru magatte ne.*

‘the handlebars became bent, you know.’

*Complicating action*

*odeko no byooin itte ne,*

‘(I) went to the forehead hospital, you know,’

*Complicating action*

*ano ne, byooin itte ne,*

‘um, you know, (I) went to the hospital, you know,’

*Complicating action*

*sorede ne, koko maite moratte ne, odeko.*

‘then, you know, (I) had this (part) bandaged, you know, the forehead.’

*Complicating action*

*sorede ne, oziityan to obaatyan ga ne, arare o motte kite,*

‘then, you know, my grandpa and grandma, you know, brought rice cake cubes,’

*Complicating action*

*tabete,*

‘(I) ate (those cakes),’

*Evaluation*

*sorede ne, daibu naotte kita.*

‘then, you know, (I) became very much okay.’

Appendage: Coda

*owari.*

‘that’s it.’

By the time children enter school, however, they seem to have mastered the major structural features of their language; for instance, they tell stories that conform to particular cultural schemata for storytelling (e.g., a topic-centered story, which begins with orientation, builds to a climax, then resolves the action, and ends with a coda). In fact, children seem to master nearly all of the linguistic features of the language to which they are exposed seemingly without specific instruction. Yet, in the light of becoming able to produce culture- and language-specific stories that meet cultural norms, the gradual but steady socialization process by caregivers contributes greatly to the shaping of children's narrative ability.

This chapter first presented a historical overview of the field of language studies. The late-1950s witnessed a rapid development in child language studies, from the behaviorist theory of language put forth by Skinner (1957) to the Chomskian revolution (Chomsky 1957), which provided language researchers with new models to explore. The chapter then discussed developments in our understanding of how children learn to talk, with particular emphasis on the roles of innate, cognitive, and social interactive factors in language development, as well as cross-cultural differences. Language production – not only at the syntactic level but also at the discourse level – is a combination of the nature of human thought and the structural properties peculiar to an individual's native language.

This chapter has particularly focused on various aspects associated with pragmatic development, such as (a) the acquisition of culturally specific rules for using speech and (b) factors influencing language development (e.g., the role of maternal input and scaffolding behavior). The chapter has emphasized that narrative styles reflect the society and culture in which they are employed. Through narrative, an individual organizes his or her experience under the constraint of sociocultural meanings; thus, narrative can be viewed as a microcosm of the individual mind, but more than that, it reflects the larger social world. As Gee (1985: 11) puts it, "Just as the common core of human language is expressed differently in different languages, so the common core of communicative style is expressed differently in different cultures." Along the same lines, Cazden (1988: 24) claims that while "narratives are a universal meaning-making strategy, there is no one way of transforming experience into a story." This trend in thinking is also advanced by Bruner (1990), who argues that meaning creation is tightly yoked to a specific style of cultural representation.

Furthermore, the development of children's personal narratives reflects not only their culture but also their age. As Eisenberg (1985: 177) notes, "The ability to discuss and describe past events involves a number of cognitive, conversational, and linguistic skills not necessary when talking about objects and events that are visible when the conversation is taking place." To make matters more complicated, as previous research has revealed (e.g., Berman and Slobin 1994; Nelson 1989), cognitive, linguistic, conversational, and social-interactional dimensions seem to take different courses in the process of language acquisition. Moreover, these factors

interact in a complex fashion in narrative development. From the perspective of social forces as shaping the path from becoming a native speaker to being a proficient speaker/narrator, language proficiency in general and narrative proficiency in particular involves a complex configuration of interrelated types of knowledge, such as (a) linguistic command of the full range of available expressive options, both grammatical and lexical, (b) the cognitive ability to integrate forms and options in order to meet further advanced communicative goals and discourse functions, and (c) cultural recognition of what constitute the favored options of a given speech community (Berman 2004). To better understand this complexity, when the social interactionist paradigm is applied to the study of narrative, the bulk of research in this area considers the caregiver, particularly the mother, to be the primary agent who provides a framework for the child to learn a particular narrative style.

Understanding how narrative develops is crucial because a specific narrative style not only reflects a fundamental structure that has been culturally nurtured, but also indicates a socialization process contributing to the formation of such cultural representation. Narratives enable individuals to make sense of their experiences in culturally satisfying ways. With these understandings as a basis, examining Japanese children's personal narratives in the context of mother-child interactions can offer important insights into the sociocultural basis for language acquisition.

## 3.2 Two possible conjectures: Current narrative development and its end results

In talking about what has happened, young children typically focus on events and activities, and pay minimal attention to motivational, evaluative, and other background elements. As was observed earlier in this chapter, in L1 narrative development there exists a relationship between an individual's age and the amount of background information he or she adds to the narrative; compared to adults, young children tend to emphasize a temporal sequence of action with less emphasis on non-sequential information. It is well known that younger children may employ fewer expressive options during narrative constructions because they cannot (a) conceive of the full range of encodable perspectives from a cognitive point of view, (b) fully assess the listener's point of view from a communicative point of view, or (c) apply the full range of formal devices from a linguistic point of view (Berman and Slobin 1994; Minami 1996a, 1996b, 2002). Based on these cognitive and linguistic limitations, we may be able to conjecture that even though their L2 skills are limited in terms of syntax and vocabulary, adult Japanese-language learners are capable of producing narratives that include both foreground and background information, compared to young native speakers.

Furthermore, as seen earlier in this chapter, in the monologue form of narrative, Japanese children's tendency to tell concise stories shows a remarkable contrast to

the narrative style of North American children. In the dialogue form of narrative, Japanese children have a very limited number of utterances over turns, whereas North American children have a great number of utterances over turns. If we were allowed to hypothesize that human development – language development in particular – forms a continuum from early childhood to adulthood, then, it would not be surprising at all that Japanese adults tend to tell succinct narratives, even shorter than advanced Japanese-language learners.

## 3.3 Future research

Based on Vygotskian theory (Vygotsky 1978), past research on children's narrative development has identified that children develop the ability to narrate and make sense of text in the context of their conversations with adults around them. We need to expand this line of research by including and emphasizing other factors, such as examining the effects of mothers' choices for reading to their children on the children's subsequent reading habits (e.g., Reynolds and Evans 2009). To begin with, book reading is a highly valued and widely conducted home and school practice within and across literate cultures; behind this, there exists the belief that reading to young children is beneficial. We should include the interaction between mother and child when the mother reads to her child, because shared book reading is similar to the dialogic mode of narrative telling, in which, as we have seen, interaction proceeds through a question-answer format (e.g., mothers try to extend topics by providing requests for actions, descriptions, and evaluations, to which children respond). Similar maternal narrative styles are, in fact, observed during other narrative contexts, such as picture book reading interactions (Melzi and Caspe 2005). The ability to narrate competently thus holds significant meaning for later achievements in various developmental domains including emergent literacy.

Conversely, as one facet of children's literacy development, we may assume the contribution of home literacy patterns (e.g., storytelling) to have a predictive value on children's development of narrative productions (Stavans and Goldzweig 2008). Narrative serves a variety of important functions, such as mediating interpersonal relationships, self-presentation, making sense of experiences, and, as emphasized above, serving as a transition into literacy. Through dialogic narrative discourse (including book reading), mothers' styles of interviewing children about past events not only provide a template for children's narrative form, but they also support children's literacy development. In particular, as described earlier in this chapter, the preschool and early elementary school years are a period of extremely rapid development in the acquisition of literacy-related skills; in the early elementary school years, narratives are frequently given as writing assignments, which also serve as a bridge to the stories read there. In addition to book-reading interactions with adults, future research may need to include other aspects, such as dinner-table conversations. Research may also need to examine home environments and mother-child relationships "within culture" (i.e., social class differences).

# 4 Conclusion and future perspectives

Over the last several decades, researchers investigating cognitive, social, and developmental processes have been intrigued by the reciprocal relationships between cognition and social interaction. Some of these researchers have tried to identify important theoretical and empirical questions that address interpersonal processes such as parent-child interaction. The conceptual foundation for many current studies, in fact, was laid down a long time ago by such theorists as Vygotsky (1962, 1978). Following this line of research, the present chapter has focused on narrative development in L1 Japanese within the framework of narrative as a product of socially situated cognition.

This chapter has specifically addressed interactive factors that play a role in children's developing narrative abilities. The chapter started by reviewing different approaches to language acquisition that have had influence on the study of children's narratives. Particularly the social interaction approach (Vygotsky 1978) was re-evaluated from the perspective of narrative development. There are multiple phases in the development of narrative abilities (Berman 1995). During narrative development young children acquire a narrative schema (or action structure) and, following that schema, they come to adhere to the rules for producing well-formed texts. Gradually learning how to combine mature knowledge of narrative structure and a full repertoire of linguistic devices, children begin to integrate expressive skills with well-constructed narrative production.

In monologic narrative, using Labovian methodology (Labov 1997), we particularly focused on how foreground, plot-advancing narrative events are embedded in background circumstances and affective evaluations. In dialogic narrative, we also examined maternal scaffolding activities, and, as can be seen in cross-cultural comparison of parental styles of narrative elicitation, we learned that mothers who provide heavy scaffolding of their children's output may not necessarily elicit narratives effectively. Rather, through cross-cultural comparison of mother-child interactions, the study described in this chapter provides us with different views or possibilities of a complex web of interrelations between the development of narrative competence and the realization of storytelling performance, particularly the amount of caretaker scaffolding and the quality of children's narrative productions.

This chapter has specifically identified cross-cultural differences as well as cross-cultural commonalities. It is important to recall the cross-cultural comparison of the child's utterances over turns (i.e., the number of utterances produced by the child per turn), which indicate that whereas English-speaking mothers allow their five-year-olds to take long monologic turns and give many evaluative comments, Japanese mothers, whether living in Japan or in the U.S., simultaneously pay considerable attention to their five-year-olds' narratives and facilitate frequent turn exchanges. While Japanese mothers' frequent verbal acknowledgment may also suggest Japanese

speakers' culturally preferred pattern of co-construction, we should not disregard cross-cultural commonalities in the process of narrative development. As seen in Akio's monologic narrative, because five-year-olds are capable of providing a relatively sufficient amount of foreground information (i.e., complicating actions), compared to other components such as background information, mothers do not necessarily need to elicit further contributions from their children. This is not limited to the Japanese language. When we review or conduct cross-linguistic studies (e.g., Berman and Slobin 1994; Minami 2002), we recognize that, compared to non-sequential information, the tendency to provide foreground (i.e., sequential) information such as complicating actions in narrative develops early cross-linguistically as well as cross-culturally. It is therefore possible to argue that such commonalities are inherent to the nature of general developmental patterns in narrative.

In this way, we now realize a further need to integrate manifold factors involved in developing narrative abilities within a unified, developmentally motivated framework. To conclude, personal narratives organize experience by describing the general flow of events in a person's life. While the basic narrative structure is similar across different languages, specific contents of narrative, such as how characters and events are described and evaluated, often reveal culture-specific patterns. Moreover, we should not forget the fact that young children's narratives reflect the culture-specific values and beliefs that the mothers instill through their child-rearing practices and language socialization processes. Based on the results described in this chapter, we may be allowed to claim, at least in part, that the origins of narrative style can be traced back to conversations between parents and children. In this sense, children's personal narrative techniques and possibly literacy skills in their later years are influenced and constructed through social interactions with their mothers. Finally, this chapter has demonstrated that viewing social interactions through the perspective offered by the social interaction paradigm can greatly facilitate understanding Japanese language and culture.

## References

Au, Kathryn H. 1993. *Literacy instruction in multicultural settings*. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers.

Bakeman, Roger and John M. Gottman. 1997. *Observing interaction: An introduction to sequential analysis, 2nd edn*. New York: Cambridge University Press.

Berman, Ruth A. 1995. Narrative competence and storytelling performance: How children tell stories in different contexts. *Journal of Narrative and Life History* 5(4). 285–313.

Berman, Ruth A. 2004. Between emergence and mastery: The long developmental route of language acquisition. In Ruth A. Berman (ed.), *Language development across childhood and adolescence* (Trends in Language Acquisition Research 3) 9–34. Amsterdam, The Netherlands: John Benjamins.

Berman, Ruth A. and Dan Isaac Slobin. 1994. *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bruner, Jerome. 1977. Early social interaction and language development. In H. Rudolph Schaffer (ed.), *Studies in mother-child interaction*, 271–289. London: Academic Press.

Bruner, Jerome. 1990. *Acts of meaning*. Cambridge, MA: Harvard University Press.

Cazden, Courtney B. 1988. *Classroom discourse*. Portsmouth, NH: Heinemann.

Chang, Chien-ju. 2003. Talking about the past: How do Chinese mothers elicit narratives from their young children across time. *Narrative Inquiry* 13(1). 99–126.

Chang, Chien-ju. 2006. Linking early narrative skill to later language and reading ability in Mandarin-speaking children: A longitudinal study over eight years. *Narrative Inquiry* 16(2). 275–293.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1959. A review of B. F. Skinner's *Verbal behavior. Language* 35(1). 26–58.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Dickinson, David and Alyssa McCabe. 1991. The acquisition and development of language: A social interactionist account of language and literacy development. In James F. Kavanagh (ed.), *The language continuum: From infancy to literacy*, 1–40. Parkton, MD: York Press.

Eisenberg, Ann R. 1985. Learning to describe past experiences in conversation. *Discourse Processes* 8(2). 177–204.

Ely, Richard and Allyssa McCabe. 1993. Remembered voices. *Journal of Child Language* 20(3). 671–696.

Fivush, Robyn. 1991. The social construction of parental narratives. *Merrill-Palmer Quarterly* 37(1). 59–82.

Fivush, Robyn and Fayne A. Fromhoff. 1988. Style and structure in mother-child conversations about the past. *Discourse Processes* 11(3). 337–355.

Gee, James Paul. 1985. The narrativization of experience in the oral style. *Journal of Education* 167(1). 9–35.

Gee, James Paul. 1991. Memory and myth. In Allyssa McCabe and Carole Peterson (eds.), *Developing narrative structure* 1–25. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gwyn, Richard. 2000. "Really unreal": Narrative evaluation and the objectification of experience. *Narrative Inquiry* 10(2). 313–340.

Harkins, Debra A. and Sukanya Ray. 2004. An exploratory study of mother-child storytelling in East India and Northeast United States. *Narrative Inquiry* 14(2). 347–367.

Heath, Shirley Brice. 1983. *Ways with words: Language, life and work in communities and classrooms*. New York: Cambridge University Press.

Hicks, Deborah. 1994. Individual and social meanings. *Journal of Narrative and Life History* 4(3). 215–240.

Hoff-Ginsberg, Erika. 1992. How should frequency in input be measured? *First Language* 12(3). 233–244.

Hopper, Paul J. 1979. Some observations on the typology of focus and aspect in narrative language. *Studies in Language* 3(1). 37–64.

Hopper, Paul J. and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.

Hudson, Judith A. 1993. Reminiscing with mothers and others: Autobiographical memory in young two-year-olds. *Journal of Narrative and Life History* 3(1). 1–32.

Hudson, Judith A. and Lauren A. Shapiro. 1991. From knowing to telling: The development of children's scripts, stories, and personal narratives. In Allyssa McCabe and Carole Peterson (eds.), *Developing narrative structure*, 89–136. Hillsdale, NJ: Lawrence Erlbaum Associates.

Karmiloff-Smith, Annette. 1986. Some fundamental aspects of language development after age 5. In Paul Fletcher and Michael Garman (eds.), *Language acquisition*, 2nd edn., 455–474. New York: Cambridge University Press.

Labov, William. 1972. *Language in the inner city: Studies in the Black English vernacular*. Philadelphia, PA: University of Pennsylvania Press.
Labov, William. 1997. Some further steps in narrative analysis. *Journal of Narrative and Life History* 7(1–4). 395–415.
Labov, William. 2006. Narrative pre-construction. *Narrative Inquiry* 16(1). 37–45.
Labov, William and Joshua Waletzky. 1967. Narrative analysis: Oral versions of personal experience. In June Helm (ed.), *Essays on the verbal and visual arts: Proceedings of the 1966 annual spring meeting of the American Ethnological Society*, 12–44. Seattle, WA: University of Washington Press.
Landis, J. Richard and Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1). 159–174.
MacWhinney, Brian (ed.). 1999. *The emergence of language*. Mahwah, NJ: Lawrence Erlbaum Associates.
MacWhinney, Brian. 2000. *The CHILDES project, volume I: Tools for analyzing talk*, 3rd edn. Mahwah, NJ: Lawrence Erlbaum Associates.
Maynard, Senko K. 1993. *Discourse modality: Subjectivity, emotion and voice in the Japanese language*. Amsterdam, The Netherlands: John Benjamins.
McCabe, Allyssa. 1991. Editorial. *Journal of Narrative and Life History* 1(1). 1–2.
McCabe, Allyssa. 1993. Sentences combined: Text and discourse. In Jean Berko Gleason and Nan Bernstein Ratner (eds.), *Psycholinguistics*, 2nd edn., 275–308. Fort Worth, TX: Harcourt Brace College Publishers.
McCabe, Allyssa and Lynn S. Bliss. 2003. *Patterns of narrative discourse: A multicultural, life span approach*. Boston, MA: Allyn & Bacon.
McCabe, Allyssa and Carole Peterson. 1991. Getting the story: Parental styles of narrative elicitation and developing narrative skills. In Allyssa McCabe and Carole Peterson (eds.), *Developing narrative structure*, 217–253. Hillsdale, NJ: Lawrence Erlbaum Associates.
Melzi, Gigliana and Margaret Caspe. 2005. Variations in maternal narrative styles during book reading interactions. *Narrative Inquiry* 15(1). 101–125.
Michaels, Sarah. 1991. The dismantling of narrative. In Allyssa McCabe and Carole Peterson (eds.), *Developing narrative structure*, 303–351. Hillsdale, NJ: Lawrence Erlbaum Associates.
Miller, Peggy J. and Linda L. Sperry. 1988. Early talk about the past: The origins of conversational stories of personal experience. *Journal of Child Language* 15(2). 293–315.
Minami, Masahiko. 1996a. Japanese preschool children's personal narratives. *First Language* 16(3). 339–363.
Minami, Masahiko. 1996b. Japanese preschool children's and adults' narrative discourse competence and narrative structure. *Journal of Narrative and Life History* 6(4). 349–373.
Minami, Masahiko. 2001. Maternal styles of narrative elicitation and the development of children's narrative skill: A study on parental scaffolding. *Narrative Inquiry* 11(1). 55–80.
Minami, Masahiko. 2002. *Culture-specific language styles: The development of oral narrative and literacy*. Clevedon, England: Multilingual Matters.
Minami, Masahiko and Allyssa McCabe. 1995. Rice balls and bear hunts: Japanese and North American family narrative patterns. *Journal of Child Language* 22(3). 423–445.
Modell, Judith and Charlee Brodsky. 1994. Envisioning homestead: Using photographs in interviewing (Homestead, Pennsylvania). In Eva M. McMahan and Kim Lacy Rogers (eds.), *Interactive oral history interviewing*, 141–161. Hillsdale, NJ: Lawrence Erlbaum Associates.
Nelson, Katherine (ed.). 1986. *Event knowledge: Structure and function in development*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Nelson, Katherine. 1989. Introduction. In Katherine Nelson (ed.), *Narratives from the crib*, 1–23. Cambridge, MA: Harvard University Press.

Peterson, Carole. 2004. Mothers, fathers, and gender: Parental narratives about children. *Narrative Inquiry* 14(2). 323–346.

Peterson, Carole and Allyssa McCabe. 1983. *Developmental psycholinguistics: Three ways of looking at a child's narrative*. New York: Plenum.

Peterson, Carole and Allyssa McCabe. 1992. Parental styles of narrative elicitation: Effect on children's narrative structure and content. *First Language* 12(3). 299–321.

Philips, Susan Urmston. 1982. *The invisible culture: Communication in classroom and community on the Warm Springs Indian Reservation*. New York: Longman.

Piaget, Jean. 1959 [1926]. *The language and thought of the child*. London: Routledge & Kegan Paul.

Preece, Alison. 1987. The range of narrative forms conversationally produced by young children. *Journal of Child Language* 14(2). 353–373.

Reynolds, Kailey Pearl and Mary Ann Evans. 2009. Narrative performance and parental scaffolding of shy and nonshy children. *Applied Psycholinguistics* 30(2). 363–384.

Sachs, Jacqueline. 1979. Topic selection in parent-child discourse. *Discourse Processes* 2(2). 145–153.

Sachs, Jacqueline. 1982. Talking about the there and then: The emergence of displaced reference in parent-child discourse. In Keith E. Nelson (ed.), *Children's language, volume 4*, 1–28. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schieffelin, Bambi B. and Elinor Ochs (eds.). 1986. *Language socialization across cultures*. New York: Cambridge University Press.

Shulova-Piryatinsky, Irene and Debra A. Harkins. 2009. Narrative discourse of native and immigrant Russian-speaking mother-child dyads. *Narrative Inquiry* 19(2). 328–355

Skinner, Burrhus Frederic. 1957. *Verbal behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Snow, Catherine E. and Charles A. Ferguson (eds.). 1977. *Talking to children: Language input and acquisition*. New York: Cambridge University Press.

Snow, Catherine E. and David K. Dickinson. 1991. Skills that aren't basic in a new conception of literacy. In Edward M. Jennings and Alan C. Purves (eds.), *Literate systems and individual lives: Perspectives on literacy and schooling*, 179–218. Albany, NY: SUNY Press.

Stavans, Anat and Gil Goldzweig. 2008. Parent-child-adult storytelling: Commonalities, differences and interrelations. *Narrative Inquiry* 18(2). 230–257.

Sulzby, Elizabeth and Liliana Barro Zecker. 1991. The oral monologue as a form of emergent reading. In Allyssa McCabe and Carole Peterson (eds.), *Developing narrative structure*, 175–213. Hillsdale, NJ: Lawrence Erlbaum Associates.

Tabors, Patton O., Catherine E. Snow and David K. Dickinson. 2001. Homes and schools together: Supporting language and literacy development. In David K. Dickinson and Patton O. Tabors (eds.), *Beginning literacy with language: Young children learning at home and school*, 313–334. Baltimore, MD: Brookes Publishing.

Templin, Mildred C. 1957. *Certain language skills in children*. Minneapolis: University of Minnesota Press.

Vygotsky, Lev S. 1962. *Thought and language*. Cambridge, MA: MIT Press.

Vygotsky, Lev S. 1978. *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Watson, John B. 1913. Psychology as the behaviorist views it. *Psychological Review* 20(2). 158–177.

Watson, John B. 1924. *Behaviorism*. Chicago: University of Chicago Press.

Wortham, Stanton. 2000. Interactional positioning and narrative self-construction. *Narrative Inquiry* 10(1). 157–184.

Yasuhiro Shirai

# 7 L2 acquisition of Japanese

## 1 Introduction

It is well known that a theory of language or how language is processed and acquired is primarily based on the facts and observations in a small number of European languages, predominantly English, French, German, Dutch, Spanish, etc., which is often referred to as European language bias. This is also true in the field of second language acquisition (SLA). Most of the studies cited in standard SLA textbooks as the basis for current theories are based on data from the acquisition of English and a few European languages.

Japanese is one of the few non-European languages that have made a substantive contribution to SLA theories. In this chapter, I will review major research in the acquisition of Japanese and evaluate how the acquisition of Japanese has contributed to SLA theories. To achieve this goal, I have chosen to do a systematic survey of the literature using citation count. Citation count, though not without problems, at least gives us objective criteria of how a particular study has impacted the field of SLA and beyond.

Mori and Mori (2011) present a comprehensive review of recent research (2001–2010) on L2 acquisition and instruction of Japanese in the journal *Language Teaching*, in which they claim:

> . . . empirical research on L2 Japanese (or JSL) learning and instruction is expanding, both in quality and quantity, resulting in an increasing number of publications in journals, books, doctoral dissertations and conference proceedings specializing in L2 Japanese learning and instruction, as well as in applied linguistics in general.

Although they did not support this claim with data, I tend to agree with their observation.[1] What is important, at the same time, is whether the increase of research on Japanese has impacted the field. It is possible that research that has increased in quantity (or even quality) has not contributed much to our understanding of psycholinguistic issues related to L2 acquisition in Japanese or second/foreign language in general. Thus, in this chapter I approach this impact issue using citation records, namely Google Scholar and Web of Sciences (Thomson Citation Index).

**1** Their review is oriented more toward the audience in language teaching. For more pedagogy oriented issues in Japanese, see Minami's *Handbook of Japanese Applied Linguistics* in this series.

# 2 Method

Two steps are taken in determining citation count. The first step, using Google Scholar, identifies the research studies that have been cited widely. The second step, using the Citation Index (Thomson Learning), identifies actual citation counts which were used to rank the studies here that have impacted the field.[2]

## 2.1 Identification of research studies (Google Scholar)

I have decided to use Google Scholar to identify works on L2 Japanese acquisition primarily because of its wide coverage. Google Scholar covers a wide range of materials, not just works published in journals, but also unpublished manuscripts, handouts, powerpoint files, as long as they are housed on a server of an academic organization, or otherwise deemed scholarly. Thus it will be more useful as an initial screening of works to be entered into the database of L2 Japanese research.

I used the following key words in locating cited works on L2 Japanese:

(1) a. Japanese "second language" acquisition
b. "acquisition of Japanese" "second language" L2
c. "learners of Japanese"

Quotation marks, as is well known, only take exact strings rather than 'second' and 'language' appearing separately, as in "second son's language", thus minimizing irrelevant hits. Output of all these searches were examined up to page 30 (i.e. 300 studies) and works that involved acquisition of Japanese with 10 or more citations were extracted. This yielded 146 items. The searches were done on January 1 and 2 in 2013.

## 2.2 Ranking of citations (Citation Index)

The next step is ranking the works that had high impact based on the Thomson Citation Index. I chose to use the Citation Index rather than Google Scholar for this second step for two reasons. First, Google Scholar involves many errors. It sometimes counts erroneous citations. For example, as I was going through recent citations of one of my articles, I noticed there were several irrelevant citations being counted.

---

**2** It should also be noted that I have checked Google Scholar for works published in Japanese, using Japanese key words. However, the citations of empirical research on L2 Japanese were so low (max 13 citations) that it was obvious that none of them would make the top 30. However, citation research in Japanese is certainly called for as a future project.

Also one of my articles did not appear in citations for some years even though it was cited in various articles and books. Therefore, its coverage is wide but its accuracy is not reliable. Second, it includes anything that is on the server of an academic institution and therefore anything can be part of citation count, as noted above. Thus, the number of citations on Google Scholar, though useful, may not be as trustworthy as we might think.

The Citation Index is more selective. Its compilers include almost all major journals, including journals in non-English languages (e.g., Japanese), and have certain criteria in determining whether to include a journal in the database, whose selection is updated periodically. In the area of second language studies, for example, some journals which were not previously included (i.e., *Second Language Research, Studies in Second Language Acquisition*, and *Bilingualism: Language and Cognition*) were included in their database after 2000. The Citation Index also includes some conference proceedings, but mostly its citation counts come from the number of citations in academic journals. This means if a work is cited in a book, it will not usually be counted, whereas a book that is cited in a journal will be counted, as we shall see later.

As noted above, although Google Scholar was chosen for the first step to maximize the coverage, its citation count may not be reliable because of the nature of citing materials and its frequent errors. Therefore, the Citation Index entries were chosen for ranking the impact of the cited research. The Citation Index in Web of Science has been recognized in different parts of the world as a reliable measure of a researchers' standing in the field. They have been used in tenure and promotion cases in some disciplines in North America, and have been used recently in South Korea, Hong Kong and Taiwan for various measures of academic achievements. In People's Republic of China, publications in the journals listed in the Citation Index result in monetary reward. Although it is not without controversy, the use of the Citation Index has already been an established practice in measuring academic impact.

Citation counts were checked through the University of Pittsburgh library, which only includes journals published from 1990 and later. This was deemed acceptable because on their web page, Web of Science notes "Citing Article counts are for all databases and all years, not just for your current database and year limits." I used Arts & Humanities Citation Index and Social Science Citation Index. From my past experience, these two will cover the most relevant studies in our field even without using Science Citation Index, which yields many irrelevant items.

One might wonder whether simple citation count is a valid measure of academic impact because the older the work is, the better the chances are that the work is cited – i.e., simply more opportunities. This is probably true to a certain degree. However, there are other factors to be considered. First, a paper is most frequently cited within a few years after its publication and that is why Impact Factors, a measure of a journals' impact on the field, are calculated based on citations of the two

years prior to the current year. Also, there are more publications included in the Citation Index each year and therefore are more chances of a recent work being cited. Thus, I have just counted simple citation numbers, knowing it is not a perfect measure of a work's impact.

Of 146 works in the database, I have excluded a small number of works which did not collect data that involved learning of Japanese (e.g., a study on teacher cognition), and out of the top 60 in the original database, I came up with 30 of the most highly cited works below.

# 3 Results

Table 1 lists the thirty works on L2 Japanese that are most frequently cited in the Thomson Citation Index.

As noted above, the ranking is determined purely by citation counts. In the case of a tie, those with a higher count on Google Scholar are placed higher, and if that cannot break the tie, newer works are ranked higher. However, we should keep in mind that in fact the difference of one or two citations does not mean much.

Still the citation counts of these works, especially those in the top 10, are quite impressive. To put these numbers into perspective, I checked *Language Learning*, a premier applied linguistics journal focusing mostly on second language acquisition, which has a long history, for its most highly cited articles: the most highly cited was 229 (Norris and Ortega 2000), followed by 192 (Pica 1994) and 121 (Truscott 1991), and the 10th was 85 (Macintyre and Gardner 1989). The two top articles among Japanese L2 research (127 and 104) fair well with these citation records, indicating that JSL (Japanese as a second language) research has impacted the field to some extent. In Table 1, I will make some observations about the results.

## 3.1 Social turn?

In the field of SLA, a "social turn" has been noted as a trend since the 1990's. The idea is that mainstream SLA views a learner not as a social being but as a cognitive individual devoid of social context (e.g. Firth and Wagner 1997). The new, more socially oriented approaches to SLA have been gaining ground ever since, and these include language socialization (Ochs 1988), Vygotskyan socio-cultural theory (Lantolf and Pavlenko 1995), Conversation Analysis (Sachs, Shegloff and Jefferson 1974). Of the research in the top 30, Ohta's work (1995, 1999, 2000, 2001) is framed in language socialization and sociocultural theory, whereby she analyzed Japanese L2 classroom discourse longitudinally. Siegal's (1995, 1996) ethnographic study of a woman from New Zealand learning Japanese in Japan also falls in this category.

**Table 1:** Highly cited works on the acquisition of Japanese as a second language based on Thomson citation index (Google scholar count in parenthesis)

| Author (year) | type of publication | citation | (Google count) |
|---|---|---|---|
| 1) Ohta (2001) | book | 127 | (371) |
| 2) Long, Inagaki & Ortega (1998) | *Modern Language Journal* | 104 | (329) |
| 3) Aida (1994) | *Modern Language Journal* | 67 | (470) |
| 4) Siegal (1996) | *Applied Linguistics* | 61 | (248) |
| 5) Iwashita (2003) | *Studies in Second Language Acquisition* | 51 | (200) |
| 6) Mori, J. (2002) | *Applied Linguistics* | 49 | (172) |
| 7) Toyoda & Harrison (2002) | *Language Learning and Technology* | 48 | (160) |
| 8) Ohta (1995) | *Issues in Applied Linguistics* | 39 | (170) |
| 9) Kanagy (1999) | *Journal of Pragmatics* | 38 | (78) |
| 10) Loschy (1994) | *Studies in Second Language Acquisition* | 37 | (180) |
| 11) Ohta (2000) | Lantolf (ed.) | 36 | (240) |
| 12) Saito, Garza & Horwitz (1999) | *Modern Language Journal* | 35 | (258) |
| 13) Nagata (1993) | *Modern Language Journal* | 35 | (148) |
| 14) Kitade (2000) | *Computer Assisted Language Learning* | 33 | (181) |
| 15) Ohta (1999) | *Journal of Pragmatics* | 32 | (93) |
| 16) Koda (1989) | *Foreign Language Annals* | 29 | (114) |
| 17) Brown (1995) | *Language Testing* | 25 | (142) |
| 18) Shirai & Kurono (1998) | *Language Learning* | 25 | (97) |
| 19) Dewey (2004) | *Studies in Second Language Acquisition* | 25 | (63) |
| 20) Sasaki (1994) | *Studies in Second Language Acquisition* | 25 | (63) |
| 21) Siegal (1995) | Freed (Ed.) | 24 | (79) |
| 22) Hirata (2004) | *Journal of the Acoustical Society of America* | 24 | (40) |
| 23) Ishida (2004) | *Language Learning* | 23 | (63) |
| 24) Mori, Y. (1999) | *Language Learning* | 22 | (145) |
| 25) Chikamatsu (1996) | *Studies in Second Language Acquisition* | 22 | (74) |
| 26) Iwashita (2001) | *System* | 22 | (55) |
| 27) White (1995) | *System* | 21 | (138) |
| 28) Sasaki (1991) | *Applied Psycholinguistics* | 21 | (68) |
| 29) Inagaki (2001) | *Studies in Second Language Acquisition* | 21 | (60) |
| 30) Kondo-Brown (2005) | *Modern Language Journal* | 21 | (57) |

Kanagy's (1999) work on a Japanese immersion school is based on language socialization theory, and J. Mori's (2002) work analyzed L2 Japanese discourse using a conversational analytic method. The fact that most of these studies are in the top 15 (rather than 16 to 30) shows that a social turn, shift in research emphasis from cognitive research to socially-oriented research, has occurred in Japanese SLA.

## 3.2 Input-interaction research

This line of research is considered a more traditional mainstream cognitive-interactionist approach to SLA that is solidly among the most influential within the SLA research paradigms. It originated in Hatch's (1978) discourse theory, and was devel-

oped by Long (1980, 1996) among others. It typically analyzes classroom interaction quantitatively to investigate how (simplified/modified) input, interaction, and feedback impact second language development.

Four experimental studies are included in the top 30 studies within this framework. Long, Inagaki and Ortega (1998) investigated the effect of recast (implicit negative feedback on learner errors) in Japanese and Spanish, while Iwashita (2003) investigated the effect of recast and positive evidence (i.e. input) in language development. For example, when a learner says (2a) when in fact (2b) is the correct form, the native interlocutor often provides the correct form in the form of recast.

(2) a. *Hanasimasu.* (simple nonpast tense)
speak POL NONPST
'He will speak.'

b. *Hanasiteimasu.* (imperfecive aspect)
speak PROG POL NONPST
'He is speaking.'

Iwashita (2001) tested the effect of different types of proficiency grouping (High-High, Low-Low, and High-Low) on language development. Loschky (1995) also investigated the effect of modified input and interaction on learning vocabulary and grammar. Ishida (2004) investigated the effect of recast on the development of the aspect marker *–te i-(ru)*, as in (2) above. It is noteworthy that most of these studies were conducted at the University of Hawai'i, where Long was teaching at the time.

## 3.3 Computer-mediated instruction

Given the increasing use of computers in L2 classrooms, many of the highly cited studies involve CALL (Computer-assisted language learning). Both Kitade (2000) and Toyoda and Harrison (2002) analyzed internet chat (NS-NNS interaction). Kitade (2000) framed her analysis within the sociocultural theory framework, while Toyoda and Harrison focused on the 'negotiation of meaning' with a more traditional interactionist perspective. Both argue that internet chat is useful for language development.

Nagata (1993) conducted an experimental study which evaluated the effectiveness of different types of computer feedback on grammar: traditional feedback, which only tells the students what is wrong with their answers, vs. intelligent feedback[3], which also tells them why their responses are wrong. After six training sessions focusing on passive voice, it was shown that intelligent feedback was superior, and in particular effective for improvement in the accurate use of particles.

**3** Intelligent feedback provides information about why the learner's response is incorrect. For example, Nagata (1993: 335) states: "In your sentence, GAKUSEE is the 'subject' of the passive (the one that is affected by the action), but it should be the 'agent' of the passive (the one who performs the action and affects the subject). Use the particle NI to mark it."

## 3.4 Research on individual differences (ID)

Research on individual differences in the list included two studies on foreign language anxiety (Aida 1994; Saito, Garza and Horwitz 1999), and two studies on learner beliefs (Mori 2002 and White 1995).

Aida (1994) found that foreign language anxiety, which consists of presentation anxiety, test anxiety, and fear of negative evaluation (Horwitz, Horwizt and Cope, 1986) which has been found in the learning of European languages in the US, also plays a negative role in the learning of Japanese as a foreign language (JFL). Saito, Garza and Horwitz (1999) identified a specific type of anxiety called "foreign language reading anxiety" by comparing L1 English learners of French, Russian, and Japanese.

Y. Mori (1999) investigated the structure of learner beliefs and its relationship to their achievements in L2 learning through a questionnaire administered to learners of Japanese, and proposed six distinct factors (e.g., Avoid Ambiguity, Reliance on L1, etc.) in L2 learner beliefs based on factor analysis. White (1995) conducted a longitudinal study on learner beliefs in the context of a self-instructed university distance learning course, which investigated learners of Spanish and Japanese in New Zealand via questionnaires and interviews.

## 3.5 Other topics

There were a couple of studies which did not fit in any of the headings discussed above. They included: a study on heritage language learning by Kondo-Brown (2005), a study on study abroad by Dewey (2005), and a language testing study by Brown (1995). Kondo-Brown compared the proficiencies of non-heritage JFL learners and Japanese as heritage language (JHL) learners. She found that a particular type of JHL learner, namely learners who have a Japanese speaker as a parent, have a more superior ability than non-heritage JFL learners, but heritage learners who do not have a Japanese-speaking parent have a very similar ability to that of JFL learners.

Brown (1995) compared how the oral Japanese Language Test for Tour Guides in Australia is rated by assessors of two different backgrounds (Japanese language teachers and tour guides), and found that rater background did not significantly change their overall rating patterns, although there were some differences.

Dewey (2005) compared changes in reading comprehension ability in two settings: domestic intensive immersion (Middlebury College, VT, USA) and study abroad (Kyoto). Here again there was mostly no significant difference between the two programs.

## 3.6 Japanese psycholinguistics

In this last section of the review, I will focus on the studies that are squarely in Japanese psycholinguistics, the focus of this volume. Here, I consider (a) studies that address the issue of how Japanese linguistic structures are acquired by L2 learners;

and (b) studies in which understanding of how the properties of a particular Japanese structure are acquired contributes to general psycholinguistic issues, not just Japanese psycholinguistics.

The studies reviewed so far address general issues of second language acquisition, which happen to use Japanese as a target language. For example, Loschky's (1995) study could have easily been done on Spanish with the same design and led to the same theoretical contribution to SLA. Thus, although it makes a contribution to psycholinguistics (i.e., SLA) in general, it does not necessarily make a significant contribution to Japanese psycholinguistics.

In what follows, I review the studies of this type that were listed in the top 30, and how they contributed to issues in Japanese psycholinguistics. When I refer to the studies that are in the Google Scholar search in step 1 but not listed in the top 30, I will put their citation count in square brackets. Sections 3.6.1 to 3.6.7 review the studies that were in the top 30, and Sections 3.6.8 to 3.6.10 review the studies that were not in the top 30 but are deemed important for L2 Japanese psycholinguistics.

### 3.6.1 Transfer in reading

Koda (1989) investigated how a learner's L1 affects the development of reading comprehension (by employing a cloze task and a paragraph comprehension task) in Japanese. She compared three different L1 groups (English, Chinese, and Korean), and found that the L1 English group was at a disadvantage right from the beginning, and the gap between the English L1 group and Chinese/Korean groups, who share Chinese characters with Japanese, widened as the learners' proficiency-level went up.

Chikamatsu (1996) investigated the effects of an L1 orthographic system on L2 word recognition strategies. Lexical judgment tests using Japanese kana given to L1 English and Chinese learners revealed that Chinese learners relied more on the visual information in L2 Japanese kana words than did L1 English learners and that L1 English learners utilized the phonological information in Japanese kana words more than did Chinese learners, which suggests that native speakers of English and Chinese utilize different word recognition strategies (i.e. different levels of reliance on phonological vs. visual cues) due to L1 orthographic characteristics.

Both studies, which investigated reading of Japanese by L2 learners having different L1 orthographic systems, show that language transfer is a significant predictor of L2 reading, which had informally been pointed out but had not been shown with solid empirical evidence.

### 3.6.2 The Aspect Hypothesis

Shirai and Kurono (1998) tested the Aspect Hypothesis (Andersen and Shirai 1994; see also Gabriele and Hughes in this volume), which predicts that there is a strong

correlation between tense-aspect marking and lexical aspect; namely, between telic verbs (Vendler's 1957 achievement and accomplishment verbs) and past/perfective marking, and between progressive marking and activity verbs, which had been found in European languages such as Spanish (e.g., Andersen 1991), French and English (e.g., Bardovi-Harlig and Bergström 1996). Shirai and Kurono confirmed the same trend in two experiments with an oral interview and a grammaticality judgment task, and argued for the universal status of the Aspect Hypothesis. For example, learners' use of past tense form *–ta* was strongly associated with achievement verbs (e.g., *otiru* 'drop', *sinu* 'die'), while imperfective *–te i-(ru)* was associated with activity verbs (e.g., *utau* 'sing', *hanasu* 'talk').

Ishida (2004), reviewed above in the section of input/interaction, also addressed the same issue. She focused on *–te i-(ru)*, the imperfective aspect marker, which denotes not only progressive meaning but also resultative meaning, as in (3).

(3) *Ken wa* *sin de iru.*
Ken TOP die RES NONPST
'Ken is dead.'

Her students found resultative meaning (obtained with achievement verbs) easier than progressive meaning (obtained with activity verbs). This was attributed to input frequency, i.e. the students were exposed to resultative meaning of *–te i-(ru)*, long before they encounter the progressive use of the aspect marker, thus underscoring the importance of distributional bias as an important contributor to universal tendency observed in the domain of tense-aspect acquisition. The acquisition of tense-aspect in Japanese has since been a fruitful ground of research in SLA (see, for example, Shirai 2002 [28], Sugaya and Shirai 2007 [24], Gabriel 2009 [12]).

### 3.6.3 The Competition Model

Sasaki (1991, 1994) applied the Competition Model (Bates and MacWhinney 1987) to L2 Japanese. The Competition Model, which originated in L1 acquisition/processing research, showed through crosslinguistic research that cues for comprehension differ from language to language, some cues stronger than others. For example, word order is the strongest cue for agent identification in English, while case marking is the strongest in Japanese. Sasaki tested to see if such sensitivity to different cues acquired by learning a particular L1 is transferred in L2 processing. In a bidirectional study, Sasaki (1991) compared processing strategies of four groups: native Japanese speakers, native English speakers, intermediate or advanced JFL (L1 English) learners, and intermediate EFL (L1 Japanese) learners[4]. He found that L2 English (L1 Japanese) speakers relied more on animacy cues than native English speakers. Native Japanese

and L2 Japanese learners showed similar processing strategies as well. He suggested that this indicates that animacy-based processing strategies have universal precedence over grammatical cues at the early stages of SLA because the semantic notion of animacy is universally available for comprehension. Sasaki (1994) compared beginning and intermediate level L1 English JFL learners and L1 Japanese EFL learners for their comprehension strategies in both English and Japanese. He found that JFL learners relied more on case-marking as their proficiency increased, and that an animacy cue was prominent for JFL learners. These findings were similar to those in Sasaki (1991).

The other competition model studies include Rounds and Kanagy (1995) [23] and Sasaki (1997) [17]. For a review of competition model studies in Japanese, see Sasaki and MacWhinney (2006).

### 3.6.4 Motion verbs

Inagaki (2001) conducted a bi-directional study on the motion verb expression in English and Japanese, using a grammaticality judgment task. English is a satellite-framed language while Japanese is a verb-framed language, according to Talmy's (1985) typology. He found that motion expressions allowed in English, but not in Japanese, e.g., (4), are erroneously transferred into L2 Japanese, but are correctly accepted by Japanese learners of English since the latter can be learned from positive evidence abundantly available in the input.

(4) ?*gakkoo ni aruku.*
school to walk NONPST
'walk to school'

This study suggests that L2 learners' interlanguage tends to result in overly general grammar and thus errors like (4) above occur when the L2 structure is a subset of the corresponding L1 structure. It supports the findings in previous studies in European languages, such as dative alternation in English and French (White 1987).

**4** English and Japanese sentences used include the following examples:

(i) *The horse watches the dog* [*Uma miru inu*]
(ii) *Kisses the deer the rock* [*Kisusuru shika iwa*]
(iii) *The cow the cigarette bites* [*Ushi tabako kamu*]
Sasaki (1991:52)

### 3.6.5 Phonetics/Phonology

Hirata (2004) trained American English speakers to hear length differences of Japanese words (e.g., *ko* 'child' vs. *koo* 'this way') in two different contexts (word vs. sentence[5]). It was found that there were similar significant improvements observed in both contexts regarding their ability to identify the number of moras. However, the sentence training context was superior in that the sentence-context group did fairly well in both word- and sentence-contexts and showed improvement between pretest and posttest (word 25.5% vs. sentence 20.4%) while the word-context group had a huge difference between word-context testing (30.1%) and sentence-context testing (14.5%). In other words, word context training only helped in a similar context, while sentence context training resulted in more generalized improvement. Although it remains to be seen whether learners of Japanese can improve their listening comprehension abilities of word meanings, it appears that at the level of perception, American English speakers can learn length distinctions in Japanese.

### 3.6.6 Universal Grammar studies

Conspicuously absent in the top 30 were studies in the framework of generative grammar. They are, however, not completely absent from the longer list based on Google Scholar; in particular, Kanno's works that argue for access to Universal Grammar are cited frequently (Kanno 1997 [69], 1998a [39], 1998b [27]). Research on unaccusativity[6] is also represented well, e.g., Sorace and Shomura (2001) [40]; Hirakawa (2001) [40], and so is research on binding[7] (Thomas 1991 [80], 1995 [40]). However, none of these studies made it to top 30 because their number of citations in the Thomson Citation Index was not large. It remains to be seen whether these studies will have a stronger impact in the field in the future. (See Nakayama and

---

**5** The learners' task was to put a word in a sentence, as in below.

(i) ____*to itte kudasai*/ "Please say ______."
/*dewa* ______ *ni tsuite kangaemasu*/ "Then, I will think about ______ ."

**6** For example, Sorace and Shomura (2001) tested whether (and how) L2 learners (and native Japanese speakers) were sensitive to the unaccusative (e.g. *saru* 'leave') vs. unergative (e.g. *oyogu* 'swim') distinction in relation to quantifier floating (Miyagawa 1989) and Case-drop (Kageyama 1993).

**7** For example, Thomas (1991) investigated whether English-speaking learners of L2 Japanese were sensitive to binding (co-reference of pronominal forms) conditions in Japanese. The conditions differ from those in English. In Japanese, the equivalent of (i) can have two interpretations (i.e., *zibun* can refer to either *John* or *Paul*), allowing a long distance antecedent, unlike English.

(i) *Paul thinks John loves himself.*
(*Paul-wa John-ga zibun-ga sukida-to omotteiru*)

Yoshimura's chapter in this volume for a recent theoretical approach within this framework.)

### 3.6.7 Processability Theory

DiBiase and Kawaguchi (2002) [43] and Kawaguchi (2005) [19] within the Processability Thoery (Pienemann, 1989) also have frequent citations in the Google Scholar search, but not in the Thomson Citation Index. These studies tested typological validity of the Processablility Theory (PT) with Japanese (and Italian), which previously had only been tested with Germanic languages (e.g. German, English). The PT assumes that the second language production is constrained by complexity of information exchange in producing the form (based on Lexical Functional Grammar), and predicts that the developmental sequence universally follows the hierarchy shown below, with examples from Japanese structures:

(5) a. lemma access; (word)
b. the category procedure; (lexical morphemes) verbal inflection
c. the phrasal procedure; (phrasal information) V-*te*-V
d. the S-procedure; (inter-phrasal information) passive, causative, benefactive
e. the subordinate clause procedure

Thus, L2 learners of Japanese are predicted to produce words first, then lexical level morphemes (such as past tense), followed by complex verbal predicates involving phrasal level information exchange such as V-*te*-V (e.g. V–*te mi-ru* 'try V-ing' as in (6), and then those involving inter-phrasal information exchange such as causative constructions, as in (7), before they start producing complex sentences. Below are the sentences actually produced by learners.

| | | | | | | |
|---|---|---|---|---|---|---|
| (6) | *uindosaafin o* | *site* | *mitai* | *site mi* | *site* | *mimasita* |
| | windsurfing ACC | do COMP | try DESD | do COMP | do COMP | try POL PST |

'I want to try . . . tried windsurfing.'
(DiBiase and Kawaguchi 2002: 292)

| | | | | |
|---|---|---|---|---|
| (7) | *butyoo wa* | *Ruusii-san ni* | *kopii o* | *sasemasita* |
| | dept. chief TOP | Lucy Miss OBL | photocopy ACC | do CAUS POL PST |

'The department chief made Lucy make photocopies.'
(DiBiase and Kawaguchi 2002: 294)

The prediction was born out by the 3-year longitudinal data from a learner and one cross-sectional study from 9 learners. There was no case where exception to the developmental hierarchy was observed. For example, none of the learners produced (7) before producing (6). In addition to Japanese, Italian data also supported the PT.

### 3.6.8 The Noun Phrase Accessibility Hierarchy

Keenan and Comrie's (1977) Noun Phrase Accessibility Hierarchy (NPAH) is a typological generalization regarding the relativizability of NP types. The NPAH is a hierarchy of relativizability with the following implicational hierarchy: Subject (SU) > Direct Objecti (DO) > Indirect Object (IO) > Oblique (OBL) > Genitive (GEN) > Object of Comparison (OComp). If a language can relativize on one node in the hierarchy, then it can also relativize on any relative clause type to the the left of the hierarchy. For example, if a language has object relatives as in (8b), it will also have subject relatives as in (8a).

(8) a. *Ken o* *mita otoko* (Subject relative clause)
Ken ACC saw man
'the man who saw ken'.

b. *Ken ga* *mita otoko* (Object relative clause)
Ken NOM saw man
'the man who Ken saw'

The NPAH has been applied by SLA researchers (e.g. Gass 1979; Doughty 1991) to predict the level of acquisition difficulty of relative clauses and is claimed to be universal (e.g., Ellis 1994). This, however, was tested only on English and a few other European languages. Shirai (2007) tested this hypothesis with East Asian languages in a special issue of *Studies in Second Language Acquisition*. The papers in this issue found varying results. Ozeki and Shirai (2007) [19] did not support the prediction of the NPAH, while Kanno (2007) [17] did. Ozeki and Shirai found no difference in difficulty between subject relatives (e.g., (8a)) and object relatives (e.g., (8b)), contra the prediction of the NPAH, while Kanno found subject relatives to be easier to comprehend than object relatives in Japanese. Further studies are needed to test the crosslinguistic validity of the application of the NPAH as a predictor of L2 relative clause acquisition. (See Sawasaki and Kashiwagi-Wood's chapter in this volume for L2 learners' relative clause comprehension.)

# 4 Conclusion and future directions

This chapter reviewed studies on L2 acquisition of Japanese which have had an impact on the field by systematic analysis of citation count in Google Scholar and Thomson Citation Index. Highly cited articles represent a wide array of L2 Japanese research, but mainly from two strands: Socially oriented studies (e.g. sociocultural theory; language socialization, Conversational Analysis) and cognitive-interactionist research (input, interaction, feedback). These are followed by CALL studies. These

studies may have received high citations partly because of their broader relevance – to second/foreign language classroom teaching (i.e., pedagogy). On the other hand, more or less purely theoretical research, such as the ones grouped above as "Japanese Psycholinguistics", tends not to get as many citations.

This may in fact reflect the trend in the field of SLA in general. In the 1980s, the field of SLA, which historically has had a close connection with L2 teaching, at one point tried to move away from pedagogy to establish itself as a purely theoretical discipline (e.g., Gass 1989). However, the field has since seemingly realized its importance as an applied field of inquiry (Gass and Selinker 2001). This trend may be apparent in the field of Japanese SLA research. (See Minami's Applied Linguistics Volume.) That does not mean, of course, that we should not continue more theoretically motivated research.

What is missing from the above studies are neuroscientific studies. However, some are already emerging. Jeong et al.'s (2007) [21] fMRI study tested the cortical activation of L1 Korean, L2 English and L3 Japanese speakers, and found that the activation pattern in a listening comprehension task was similar between Japanese and Korean, but not English, suggesting the influence of language distance in neuronal activities. Nakada, Fujii and Kwee (2001) [48] showed that neural activation patterns of Japanese and English native speakers did not change when reading[8] their L1 or L2, but did differ from each other, on activation of the LG (lingual gyrus) being much higher for native English participants than Japanese participants, when reading their L1 or L2. This suggests that the L1 reading process is transferred to L2 reading.[9] Also conspicuously missing are the standard sentence processing studies often seen in adult processing of Japanese (e.g., Kamide and Mitchell 1997) as well as in L2 English (e.g., Juffs and Harrington 1995; Clahsen and Felser 2006). In particular, the head-final nature of the Japanese language provides us with many interesting grounds for testing theories developed in head-initial European languages (see Sawasaki and Kashiwagi-Wood's chapter in this volume for L2 Japanese processing research). These studies will further make Japanese SLA a theoretically important subfield of psycholinguistics.

**8** Reading passages were taken from standard proficiency tests of respective languages (e.g., TOEFL).

**9** This study was not identified in the original Google Scholar search. When I did an additional search looking for processing research in L2 Japanese by the key words [Japanese "second language" processing], it was identified. This study had 25 citations in the Citation Index, which would have been in the top 20.

## Acknowledgments

This paper was written while the author was affiliated with the National Institute for Japanese Language and Linguistics as an invited scholar.

## References

Aida, Yukie. 1994. Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal* 78(2). 155–168.

Andersen, Roger W. 1991. Developmental sequences: The emergence of aspect marking in second language acquisition. In Thom Huebner and Charles A Ferguson (eds.), *Crosscurrents in second language acquisition and linguistic theories*, 305–324. Amsterdam: John Benjamins.

Andersen, Roger W. and Yasuhiro Shirai. 1994. Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition* 16(2). 133–156.

Bardovi-Harlig, Kathleen and Anna Bergström. 1996. Acquisition of tense and aspect in second language and foreign language learning: Learner narratives in ESL and FFL. *Canadian Modern Language Review* 52(2). 308–330.

Bates, Elizabeth and Brian MacWhinney. 1987. Competition, variation, and language learning. In Brian MacWhinney (ed.), *Mechanisms of language acquisition*, 157–193. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, Anne. 1995. The effect of rater variables in the development of an occupation-specific language performance test. *Language Testing* 12(1). 1–15.

Chikamatsu, Nobuko. 1996. The effects of L1 orthography on L2 word recognition: A study of American and Chinese learners of Japanese. *Studies in Second Language Acquisition* 18 (4). 403–432.

Clahsen, Harald and Claudia Felser. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27(1). 3–42.

Di Biase, Bruno and Satomi Kawaguchi. 2002. Exploring the typological plausibility of Processability Theory: Language development in Italian second language and Japanese second language. *Second Language Research* 18(3). 274–302.

Doughty, Catherine. 1991. Second language instruction does make a difference. *Studies in Second Language Acquisition* 13(4). 431–469.

Ellis, Rod. 1994. *The study of second language acquisition*. Oxford, UK: Oxford University Press.

Firth, Alan and Johannes Wagner. 1997. On discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal* 81(3). 285–300.

Gabriele, Alison. 2009. Transfer and transition in the SLA of aspect. *Studies in Second Language Acquisition* 31(3). 371–402.

Gabriele, Alison and Mamori Sugita Hughes. 2015. Tense and aspect in Japanese as a second language. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Gass, Susan. 1979. Language transfer and universal grammatical relations. *Language Learning* 29(2). 327–344.

Gass, Susan M. 1989. Language universals and second-language acquisition. *Language Learning* 39(4). 497–534.

Gass, Susan M. and Larry Selinker. 2001. *Second language acquisition: An introductory course.* 2nd edn. Mahwah, NJ: Lawrence Erlbaum Associates.

Hatch, Evelyn M. 1978. Discourse analysis and second language acquisition. In Evelyn M. Hatch (ed.) *Second language acquisition: A book of readings*, 401–435. Rowley, MA: Newbury House.
Hirakawa, Makiko. 2001. L2 acquisition of Japanese unaccusative verbs. *Studies in Second Language Acquisition* 23(2). 221–245.
Hirata, Yukari. 2004. Training native English speakers to perceive Japanese length contrasts in word versus sentence contexts. *The Journal of the Acoustical Society of America* 116 (4). 2384–2394.
Horwitz, Elaine K., Michael B. Horwitz, and Joann Cope. 1986. Foreign language classroom anxiety. *The Modern Language Journal* 70(2). 125–132.
Inagaki, Shunji. 2001. Motion verbs with goal PPs in the L2 acquisition of English and Japanese. *Studies in Second Language Acquisition* 23(2). 153–170.
Ishida, Midori. 2004. Effects of recasts on the acquisition of the aspectual form *-te i-(ru)* by learners of Japanese as a foreign language. *Language Learning* 54(2). 311–394.
Iwashita, Noriko. 2001. The effect of learner proficiency on interactional moves and modified output in nonnative–nonnative interaction in Japanese as a foreign language. *System* 29(2). 267–287.
Iwashita, Noriko. 2003. Negative feedback and positive evidence in task-based interaction. *Studies in Second Language Acquisition* 25(1). 1–36.
Jeong, Hyeonjeong, Motoaki Sugiura, Yuko Sassa, Tomoki Haji, Nobuo Usui, Masato Taira, Kaoru Horie, Shigeru Sato, and Ryuta Kawashima. 2007. Effect of syntactic similarity on cortical activation during second language processing: A comparison of English and Japanese among native Korean trilinguals. *Human Brain Mapping* 28(3). 194–204.
Juffs, Alan and Michael Harrington. 1995. Parsing effects in second language sentence processing. *Studies in Second Language Acquisition* 17(4). 483–516.
Kamide, Yuki and Don C. Mitchell. 1997. Relative clause attachment: Nondeterminism in Japanese parsing. *Journal of Psycholinguistic Research* 26(2). 247–254.
Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.
Kanagy, Ruth. 1999. Interactional routines as a mechanism for L2 acquisition and socialization in an immersion context. *Journal of Pragmatics* 31(11). 1467–1492.
Kanno, Kazue. 1997. The acquisition of null and overt pronominals in Japanese by English speakers. *Second Language Research* 13(3). 265–287.
Kanno, Kazue. 1998a. The stability of UG principles in second-language acquisition: Evidence from Japanese. *Linguistics* 36(6). 1125–1146.
Kanno, Kazue. 1998b. Consistency and variation in second language acquisition. *Second Language Research* 14(4). 376–388.
Kanno, Kazue. 2007. Factors affecting the processing of Japanese relative clauses by L2 learners. *Studies in Second Language Acquisition* 29(2). 197–218.
Kawaguchi, Satomi. 2005. Argument structure and syntactic development in Japanese as a second language. In Manfred Pienemann (ed.), *Cross-linguistic aspects of Processability Theory*, 253–298. Amsterdam: John Benjamins.
Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.
Kitade, Keiko. 2000. L2 learners' discourse and SLA theories in CMC: Collaborative interaction in Internet chat. *Computer Assisted Language Learning* 13(2). 143–166.
Koda, Keiko. 1989. The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals* 22(6). 529–540.
Kondo-Brown, Kimi. 2005. Differences in language skills: Heritage language learner subgroups and foreign language learners. *The Modern Language Journal* 89(4). 563–581.
Lantolf, James P. and Aneta Pavlenko. 1995. Sociocultural theory and second language acquisition. *Annual Review of Applied Linguistics* 15. 108–124.

Long, Michael H. 1980. Input, interaction and second language acquisition. Los Angeles, CA: University of California dissertation.
Long, Michael H. 1996. The role of the linguistic environment in second language acquisition. In William C. Ritchie and Tej K. Batia (ed.) *Handbook of second language acquisition*, 413–468. San Diego, CA: Academic Press.
Long, Michael H., Shunji Inagaki, and Lourdes Ortega. 1998. The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal* 82(3). 357–371.
Loschky, Lester. 1994. Comprehensible input and second language acquisition. *Studies in Second Language Acquisition* 16(3). 303–323.
Minami, Masahiko (ed.). Forthcoming. *Handbook of Japanese applied linguistics*. Boston: De Gruyter Mouton.
Miyagawa, Shigeru. 1989. *Structure and case marking in Japanese*. New York: Academic Press.
Mori, Junko. 2002. Task design, plan, and development of talk-in-interaction: An analysis of a small group activity in a Japanese language classroom. *Applied Linguistics* 23(3). 323–347.
Mori, Yoshiko. 1999. Epistemological beliefs and language learning beliefs: What do language learners believe about their learning? *Language Learning* 49(3). 377–415.
Mori, Yoshiko and Junko Mori. 2011. Review of recent research (2000–2010) on learning and instruction with specific reference to L2 Japanese. *Language Teaching* 44(4). 447–484.
Nagata, Noriko. 1993. Intelligent computer feedback for second language instruction. *The Modern Language Journal* 77(3). 330–339.
Nakada, Tsutomu, Yukihiko Fujii and Ingrid L. Kwee. 2001. Brain strategies for reading in the second language are determined by the first language. *Neuroscience Research* 40(4). 351–358.
Nakayama, Mineharu and Noriko Yoshimura. 2015. The modularity of grammar in L2 acquisition. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.
Norris, John M. and Lourdes Ortega. 2000. Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50(3). 417–528.
Ochs, Elinor. 1988. *Culture and language development: Language acquisition and language socialization in a Samoan village*. Cambridge, UK: Cambridge University Press.
Ohta, Amy Snyder. 1995. Applying sociocultural theory to an analysis of learner discourse: Learner-learner collaborative interaction in the zone of proximal development. *Issues in Applied Linguistics* 6(2). 93–121.
Ohta, Amy Snyder. 1999. Interactional routines and the socialization of interactional style in adult learners of Japanese. *Journal of Pragmatics* 31(11). 1493–1512.
Ohta, Amy Snyder. 2000. Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In James P. Lantolf (ed.), *Sociocultural theory and second language learning*, 51–78. Oxford: Oxford University Press.
Ohta, Amy Snyder. 2001. *Second language acquisition processes in the classroom: Learning Japanese*. Mahwah, NJ: Lawrence Erlbaum Associates.
Ozeki, Hiromi and Yasuhiro Shirai. 2007. Does the noun phrase accessibility hierarchy predict the difficulty order in the acquisition of Japanese relative clauses? *Studies in Second Language Acquisition* 29(2). 169–196.
Pica, Teresa. 1994. Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning* 44(3). 493–527.
Pienemann, Manfred. 1998. *Language processing and second language development: Processability theory*. Amsterdam: John Benjamins.
Rounds, Patricia L. and Ruth Kanagy. 1998. Acquiring linguistic cues to identify AGENT. *Studies in Second Language Acquisition* 20(4). 509–542.

Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4). 696–735.

Saito, Yoshiko, Thomas J. Garza, and Elaine K. Horwitz. 1999. Foreign language reading anxiety. *The Modern Language Journal* 83(2). 202–218.

Sasaki, Yoshinori. 1991. English and Japanese interlanguage comprehension strategies: An analysis based on the competition model. *Applied Psycholinguistics* 12(1). 47–73.

Sasaki, Yoshinori. 1994. Paths of processing strategy transfers in learning Japanese and English as foreign languages. *Studies in Second Language Acquisition*16(1). 43–72.

Sasaki, Yoshinori. 1997. Individual variation in a Japanese sentence comprehension task: Form, functions, and strategies. *Applied Linguistics* 18(4). 508–536.

Sasaki, Yoshinori and Brian MacWhinney. 2006. The competition model. In Mineharu Nakayama, Reiko Mazuka and Yasuhiro Shirai (eds.), *The handbook of East Asian psycholinguistics: Vol. 2, Japanese*, 307–314. Cambridge, UK: Cambridge University Press.

Sawasaki, Koichi and Akiko Kashiwagi-Wood. 2015. Issues in L2 Japanese sentence processing: similarities/differences with L1 and individual differences in working memory. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Shirai, Yasuhiro (ed). 2007. *The acquisition of relative clauses and the Noun Phrase Accessibility Hierarchy: A universal in SLA?* Special Issue of *Studies in Second Language Acquisition*, 29(2). Cambridge: Cambridge University Press.

Shirai, Yasuhiro and Atsuko Kurono. 1998. The acquisition of tense-aspect marking in Japanese as a second language. *Language Learning* 48(2). 279–244.

Siegal, Meryl. 1996. The role of learner subjectivity in second language sociolinguistic competency: Western women learning Japanese. *Applied Linguistics* 17(3). 356–382.

Siegal, Meryl. 1995. Individual differences and study abroad: Women Learning Japanese in Japan. In Barbara Freed (ed.), *Second language acquisition in a study abroad context*, 225–244. Amsterdam: John Benjamins.

Sorace, Antonella and Yoko Shomura. 2001. Lexical constraints on the acquisition of split intransitivity. *Studies in Second Language Acquisition* 23(2). 247–278.

Sugaya, Natsue and Yasuhiro Shirai. 2007. The acquisition of progressive and resultative meanings of the imperfective aspect marker by L2 learners of Japanese: Transfer, universals, or multiple factors? *Studies in Second Language Acquisition* 29(1). 1–38.

Thomas, Margaret. 1991. Universal Grammar and the interpretation of reflexives in a second language. *Language* 67(2). 211–239.

Thomas, Margaret. 1995. Acquisition of the Japanese reflexive *zibun* and movement of anaphors in Logical Form. *Second Language Research* 11(3). 206–234.

Truscott, John. 1996. The case against grammar correction in L2 writing classes. *Language learning* 46(2). 327–369.

Toyoda, Etsuko and Richard Harrison. 2002. Categorization of text chat communication between learners and native speakers of Japanese. *Language Learning & Technology* 6(1). 82–99.

Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review* 66(2). 143–160.

White, Cynthia. 1995. Autonomy and strategy use in distance foreign language learning: Research findings. *System* 23(2). 207–221.

White, Lydia. 1987. Markedness and second language acquisition. *Studies in Second Language Acquisition* 9(3). 261–286.

Mineharu Nakayama and Noriko Yoshimura

# 8 The modularity of grammar in L2 acquisition

## 1 Introduction

It has been more than half a century since a theory of Universal Grammar (UG) was proposed and various aspects of the theory have been investigated in the field of language acquisition.[1] In particular, Chomsky (1981) has had a profound influence on the development of second language (L2) acquisition theories (see White 1989, 2003b). Much of earlier L2 acquisition studies within the Principles and Parameters approach dealt with the UG accessibility (e.g., parameter-resetting approach) (for a general overview, see White 2000). For instance, the following questions were addressed: Are L1 and L2 acquisition governed by the same principles or strategies? Are grammatical principles in UG proposed in L1 acquisition fully or partially accessible during the course of L2 acquisition? Do parametric values that account for L1 variations also explain inter-language variations? What is the initial state of L2 grammar? Does it have L1 parametric values (Schwartz and Sprouse 1994) or the initial (or default) values in the grammar during the course of L1 acquisition? These questions brought certain insightful outcomes in the parameter-resetting approach, but also there remained unexplained data.

Recent developments in the Minimalist theory of grammar (Chomsky 1993, 1995) together with the interface hypothesis of L2 acquisition have made a shift in L2 acquisition research. For instance, Sorace and Filiaci (2006) report that discourse-constrained pronominal uses are delayed in L2 Italian. Advanced L2 Italian learners select *Paola* as the antecedent of *lei* in the embedded clause in (1) whereas native speakers of Italian tend to interpret *Marta* of the matrix object as its antecedent. In (2) since the complement clause precedes the matrix clause, advanced L2 Italian learners select *Paola* as the antecedent of *lei*, but native speakers of Italian select an extra-sentential antecedent.

**1** This chapter discusses L2 acquisition, in particular, adult foreign language acquisition, from a particular generative perspective. There are different views within the generative perspective and different theoretical approaches to investigating the mechanism of L2 acquisition (cf. Chomsky 2007a, b, c, 2012; Reinhart 2006). For instance, one could consider a set operation (including predication and hierarchy) and recursiveness as general cognitive primitives and our computational system constrained by (language) input, time, and working memory, which would generate language variations (e.g., word order, locality) and consider language to be learned by this computational system. This kind of cognitive nativism does not assume Universal Grammar (for different proposals, see Wolfe-Quintero 1996; O'Grady 1996; Eckman 1996; Hamilton 1996; For other theoretical approaches, see Shirai's chapter in this volume).

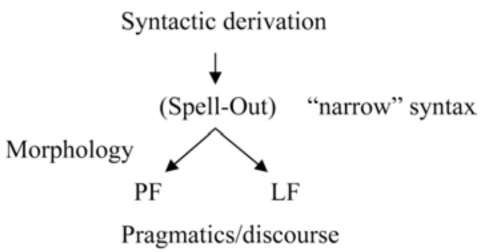
(1) *Paola*$_i$ *telefonerà a Marta*$_j$ *quando lei*$_{i/j}$ *averà tempo.*
‘Paola will telephone Marta when she will have time.’

(2) *Quando lei*$_{i/j}$ *era in vacanza, Paola*$_i$ *è andata a trovare Marta.*
‘When she was on holiday, Paola went to visit Marta.’

Since the relationship between pronouns and their antecedents involves Italian pragmatics and understanding of discourse, it is difficult even for advanced learners to acquire.[2] Although this instance is for advanced speakers, it is proposed that this interface approach can also apply to lower proficiency levels as well, as it allows us to look at the nature of the interlanguage. It uncovers the mechanism of language acquisition in a more dynamic way.

The interface theory assumes grammar modules such as lexicon, syntax, morpho-phonology, semantics-pragmatics that interact with each other (for a general review and different models within this approach, see White 2011; see also Reinhart 2006, Slobakova 2009). This approach supports the idea that syntactic properties that fall solely in a syntax module, i.e., “narrow” syntax, are acquired early while the acquisition of interface properties is delayed in L2 acquisition. Here we assume the following model.

(3)




PF is a morpho-syntax interface that may interact with morphology-phonology and articulartory-perceptual/(sensorimotor) interfaces and LF is a syntax-semantics interface that relates to conceptual-intentional interface. Core or “narrow” syntax is the domain of syntactic operations. The errors that fall between modules are more complicated as the sources of errors lie in more than one module. Keeping this general model in mind, in this chapter we will look at some L2 findings in syntax-morpho-phonology, syntax-semantics, and syntax-pragmatics interfaces as well as “narrow” syntax phenomena. In particular, we discuss the acquisition of subjects (including expletives), tense and agreement morphology, and WH-movement by Japanese EFL learners and interpretations of pronouns and reflexives in both L2 English and L2 Japanese.

**2** Similarly in L2 Japanese, it’s been shown in Nakahama (2011) that English-speaking learners of Japanese can use null pronouns, but intermediate learners use more lexical nouns (as well as overt pronouns) than those with higher proficiency. Even high proficiency speakers, however, use more full nouns in narratives than Japanese native speakers.

# 2 Morpho-syntactic interface and narrow syntax

Grammatical issues that clearly fall in the morpho-syntactic interface domain are rather difficult to observe in L2 Japanese.[3] Because of this, we will discuss the L2 acquisition of English inflectional morphology by Japanese-speaking learners of English. We compare this with the acquisition of English subjects and nominative case in this section. We look at syntax and morpho-syntax issues that are different in English and Japanese. English has the third person singular subject-verb agreement -*s*, but Japanese does not have this type of agreement.[4] Japanese allows null subjects in tensed clauses, but English does not. Suppose that the fact that learners' L1 does not have a certain grammatical phenomenon becomes a source of difficulty in L2 acquisition. Then, we would predict that Japanese-speaking learners of English would have difficulty in producing both subject-verb agreement and overt subjects.

## 2.1 Japanese EFL learners' acquisition of verbal inflection and subjects in L2 English (Syntax-morphology/phonology interface)

It has been reported in L2 acquisition studies that there is a significant discrepancy between the variable use of verbal inflection such as third person singular –*s* and

**3** L2 Japanese errors reported in Kanagy (1994) like *yasui-nai-desu* ('not cheap') and *akai-zya-arimasen* ('not red') are morphological or lexical misclassification errors (i.e., adjective vs. adjectival noun). Although the *akai-nai* ('not red'), *akai-ku-nai*, and *aka-ku-nai* sequence is observed in L1 Japanese and are morpho-syntactic errors (Sano 2002), this developmental sequence does not occur in L2, and the nature of error seems different in L1 and L2. L1 children are acquiring syntactic categories and their brains are maturing (e.g., the root infinitive stage; see Murasugi in this volume). However, adult L2 learners are both cognitively and linguistically mature and they have acquired one language already (i.e., the root infinitive stage doesn't exists in L2). Therefore, we regard these L2 errors as morphological or lexical errors. In other words, those errors in L1 and L2 may look the same, but their sources are different, and we are not certain if errors that solely fall in morpho-syntactic cases exist in L2 Japanese.

**4** In order to explain their experimental results that Japanese-speaking learners of English were insensitive to the number feature, but sensitive to the person feature in subject-verb agreement, Wakabayashi (1997), Wakabayashi et al. (2007), and Shibuya and Wakabayashi (2008) consider Japanese not to have the number feature, but to have the person feature, and point out that subject-honorification is an instance of agreement in the person feature. See also Wakabayashi and Yamazaki (2006) and Shibuya, Wakabayashi and Yamazaki-Hasegawa (2009). Since it exists in L1, they were sensitive to the feature in their experiments in English. However, Osterhout and Inoue (2007) observe N400 effects on sentences like (i) in their event-related potentials (ERP) experiment, which is a response to a more pragmatics/semantic anomalous string. Thus, subject-honorification is not syntactic agreement like English subject-verb agreement, whose violation is observed as a P600 effect.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (i) ?*Watasi* | *ga* | *zisin* | *o* | *motte* | <u>*osusume ni naru*</u> | *syoohin wa* | *kore desu.* |
| I | NOM | confidence | ACC | have | recommend HON | product TOP | this COP |

'This is a product that I confidently recommend-promoting (to you).'

regular past -*ed* and the consistent use of subjects and nominative case-marked pronouns in L2 English (White 2003a: 130; Lardiere 1998a, b, 2000; Ionin and Wexler 2002). For instance, Ladiere (1998a, b) reports an adult Chinese speaker's oral and e-mail data in which –*s* and –*ed* were correctly used only 4.5% and 34% of the time, respectively, while overt subjects and nominative case marked pronouns were used 98% and 100% of the time, respectively. A summary of previous studies, including our studies, is shown in Table 1 and some of the examples are listed in (4) and (5).

(4) *Everyone who believe it can get it.*
(Lardiere 1998a)

(5) *One time I watch this movie.*
(Ionin and Wexler 2002)

**Table 1:** Correct uses of tense and agreement morphemes, subjects, and nominative case

| Previous studies | Participant's L1 | Data Type | *-s* | *-ed* | Subject | Nominative case |
|---|---|---|---|---|---|---|
| Lardiere 1998a, b | Chinese | Oral e-mail | 4.5% | 34% | 98% | 100% |
| Haznedar 2001 | Turkish | Oral | 46.5% | 25.5% | 99% | 99.9% |
| Ionin & Wexler 2002 | Russian | Oral | 22% | 42% | 98% | – |
| Yoshimura & Nakayama 2009, 2010a | Japanese | Written Essay | 72.7% | 94.2% | 99.2% | 100% |
| Yoshimura & Nakayama 2010a | Japanese | Oral | 87.5% | 88.5% | 99.7% | 100% |

Chinese, Turkish, and Russian allow null subjects. Russian and Turkish have inflectional morphology, but Chinese does not. If L1 is the source of L2 difficulty, then it would be predicted that Turkish and Russian speakers would not have difficulty in inflectional morphology as do Chinese speakers. They all would have difficulty in the uses of overt subjects. However, these predictions are not borne out. Then why is inflectional morphology difficult to acquire? Example (4) shows the missing –*s* on the verb and (5) demonstrates the missing –*ed*. These show common variability in the domain of inflectional morphology, but are these due to surface omissions or grammatical impairment? There are plausible theoretical accounts: For instance, Eubanket et al. (1997) and Beck (1998) consider the functional features impaired in interlanguage grammar, i.e., syntactic deficits. Hawkins and Chan (1997), Hawkins and Liszka (2003), and Hawkins (2005) claim that functional categories not selected in L1 are unacquirable in L2, i.e., partial availability in adult L2 acquisition.

This is referred to as the Representational Deficit Hypothesis. On the other hand, Haznedar and Schwartz (1997) claim that there is neither syntactic deficit nor underspecification of tense/agreement in learners' grammars. Prévost and White (2000a: 108) consider "it is at the surface morphological level that inflection is assumed to be absent, rather than the abstract feature level." This is known as the Missing Surface Inflection Hypothesis (Goad and White 2004). These two hypotheses have been evaluated in Yoshimura and Nakayama (2009, 2010a).[5]

Yoshimura and Nakayama (2009, 2010a) analyzed 803 sentences in 88 written compositions collected from 44 Japanese university students studying English in study abroad contexts. Based on their scores on the Michigan Test of English Language Proficiency (maximum 100 points), a High Group made up of the 15 highest scoring students (average 78.1) and a Low Group of the 15 lowest scoring (average 52.7) were identified. A more detailed breakdown of Yoshimura and Nakayama's composition data in Table 1 is shown in Table 2 below. Three missing (or null) subjects were produced by three learners in the Low Group while two were produced by two learners in the High Group. Note that the percentage was the average percentage of the individuals' responses (e.g., percentages of the missing -*s*), not a simple percentage over the total (i.e., 19 missing -*s*/35 instances in the Low Group).

**Table 2:** Subjects in obligatory contexts, nominative case, agreement, and past tense morphology

| | 3rd Person Singular –*s* | | Past tense –*ed* | | Subject | | Nominative case | |
|---|---|---|---|---|---|---|---|---|
| | Present | Missing (%) | Present | Missing (%) | Present | Missing (%) | Present | Missing (%) |
| Low (n = 15) | 16 | 19 (41.9) | 7 | 2 (2.7) | 276 | 3 (1.1) | 121 | 0 (0) |
| High (n = 15) | 27 | 7 (12.7) | 14 | 3 (8.9) | 358 | 2 (0.6) | 205 | 0 (0) |
| Total (n = 30) | 43 | 26 (27.3) | 21 | 5 (5.8) | 634 | 5 (0.8) | 326 | 0 (0) |

The number of missing subjects was very low in both groups and there was no significant developmental difference. Since Japanese permits null subjects (pro-drophood), missing subjects would be expected in L2 English if L1 transfer occurs. Apparently, this aspect is not part of L1 transfer.[6] It is also noteworthy that Table 2 shows that both groups correctly produced nominative case marked subjects. These results indi-

**5** We explained the variable use of inflectional morphology by referring to the morphological merger (Embick and Noyer 2001; Embick and Marantz 2008) in the PF component after Spell-Out (in (3)) that the omission of inflection occurs in L2 English, unlike the feature checking before Spell-Out (Chomsky 1995).

**6** Cf. Junior high school learners in Wakabayashi (2002). See also Suda and Wakabayashi (2007). Note that this finding is different from Phinney's (1987) findings with Spanish speakers, whose native tongue also allows null subjects. Despite the difference, at least, it appears that all speakers of L1 *pro*-drop languages do not always produce null subjects in their English in the same ratio.

cate that Japanese EFL learners did not show any impairment of their L2 grammar with respect to the overtness and case-marking of subjects.[7]

As for the supplying of –*s* in obligatory contexts, the group difference was significant. As illustrated in (6) and (7), the omission errors were found in both regular and irregular forms. The variable uses of the morpheme –*s* were observed in 13 students' essays (eight students in the Low Group and five in the High Group) whereas the overuse of –*s* was infrequent (two out of 185).[8] The omission rates of the past tense morpheme were, on the other hand, relatively low in both groups, and the difference in the omission rate between the two groups was not statistically significant. Some omission examples of past tense are provided in (8).[9]

(6) **But if we use search engine on the internet, it* <u>*reduce*</u> *the time …*

(7) **This invention* <u>*have*</u> *some positive effects and negative effects.*

(8) a. **In 1903, Wight brothers* <u>*succeed*</u> *in first flight in the U.S.*
b. *… *when you* <u>*go*</u> *to NY from Tokyo, it took almost 30 days by ship.*

Although these are written data, our findings were similar to those of the oral studies. There is a significant discrepancy between the variable use of verbal inflection such as –*s* and –*ed* and the consistent use of subjects and nominative case-marked pronouns in L2 English. Assuming the grammar model in (3), functional categories are manipulated to generate a sentence structure, and the relevant structure is then sent to PF to receive morphological representation (Halle and Marantz 2000; Embick and Marantz 2008). Technically speaking, subject raising for EPP feature checking/valuation is done by virtue of merger during the course of syntactic derivations whereas abstract features associated with inflection are morphologically manifested by virtue of lexical insertion in the PF component.[10] In other words, EPP

**7** As for objects, the Low Group's missing percentage was 1.18 while that of the High Group was 1.99. Erroneous uses of accusative case were not observed in either group. Null subjects and objects are not transferred. As mentioned, this means that a L2 Root Infinitive (or optional infinitive) stage does not exist. See Yoshimura and Nakayama (2010b).

**8** Over-application requires more energy than under-application, i.e., an economic reason.

**9** Six overuse errors were found in the Low Group, one for regular and five for irregular forms, while three overuse errors were found in the High Group, one for regular and two for irregular forms. Most errors seem to have derived from the learners' difficulties with the understanding of tense as distinct from aspect in English. In addition, Yoshimura and Nakayama (2010a) report a significant difference in (singular, plural, and past tense form) errors between BE verbs and non-BE, regular verbs, which may be evidence supporting Lasnik's (1995) proposal that BE verbal forms are inserted in lexicon. At least it suggests that BE and regular verbal forms are formed differently.

**10** EPP stands for Extended Projection Principle (i.e., all sentences contain subjects). We assume this constraint must be met at PF. To be more technical here, tense/agreement features on T(ense) are checked by receiving the values of person and number features of the subject, and then, an EPP feature triggers the movement of the subject to the specifier position of TP (Tense Phrase), and nominative case is realized if T is finite (Radford 1997, 2009).

feature checking and nominative case assignment associated with subject raising are both before-spell-out operations in the syntactic component ("narrow" syntax).

On the other hand, the use of the past tense *–ed* proved to be of relatively low difficulty for most Japanese EFL learners compared to the use of the third person singular *–s* (cf. Shibuya and Wakabayashi 2008). The difference in the omission rate between *–s* and *–ed* was statistically significant. Why is the third person singular *–s* more difficult than the past tense *–ed* for the Japanese EFL learners? These two inflectional suffixes are morphologically realized in the PF component. Notice that Japanese is inflected for tense: Present/non-past tense is marked with *-u*, while past tense appears with *-ta*. That is, Japanese has the [+/-past] feature, like English, but does not have the [number] feature, unlike English. As such, if Japanese EFL learners suffer from L1 effects, they are expected to make more errors on *–s* than *–ed* because they need to learn that there is a [number] feature associated with the subject-verb agreement in English, and if it is [singular] in a present tensed finite clause, it should be spelled out by *–s*. In short, given that this "number learning" takes time and imposes a burden on Japanese EFL learners, *–s* should be more difficult than *–ed*. This explains what we saw in the learners' errors.

If this analysis is correct, the errors were not due to a prosodic problem (Prévost and White 2000a, b; Prévost 2008a) because *–s* and *–ed*, which are both on the right edge, should be equally difficult for the learners, according to the prosodic account (Goad and White 2004; White 2008). On the L1 transfer account, the learners should find the morpheme *-s* more difficult than the morpheme *–ed* because the former feature combination is not lexicalized in their L1. Our data indicated a significant difference between the omission of *–s* and *–ed* and the error ratio between *–s* and *–ed* changed over the improvement of L2 English, i.e., increased sensitivity on *–s*. This also indicates that morphological mapping improves through development and implies that the results support, to some extent, the Missing Surface Inflection Hypothesis and reject the Representational Deficit Hypothesis. As shown in Table 1, Yoshimura and Nakayama (2010a) further supports the current view that morphological mapping improves as proficiency increases.[11]

In summary, these findings were interpreted as showing that L2 English grammar of the Japanese college EFL learners: a) suffers no serious transfer from L1 prodrophood,[12] b) includes EPP feature checking and overt subject raising, and c) fails to do morphological insertion in the PF component. In other words, while the

**11** Yoshimura and Nakayama (2009, 2010a) also discuss the plural *–s* errors did not always parallel the proficiency. This is because it is heavily related to lexical acquisition (i.e., count/mass distinction).

**12** The missing objects seemed to be related to the complexity of the structure (e.g., processing). Null subjects and objects are not transferred from L1 Japanese. See Yoshimura and Nakayama (2010a).

learners could do EPP-feature checking during the syntactic merger before Spell-Out in (3), they invariably inserted the morpheme *-s* into the terminal node after Spell-Out. The learners with lower proficiency found this insertion difficult due to the absence of such morphological operation in their L1. If this analysis is on the right track, their difficulty is not due to a prosodic problem nor impaired functional categories or features in L2 grammar, but rather due to missing morphemes. This supports the Missing Surface Inflection Hypothesis (Haznedar and Schwartz 1997; Prévost and White 2000b), but rejects the Representational Deficit Hypothesis (Hawkins and Liszka 2003; Hawkins 2005).

## 2.2 Japanese learners' acquisition of expletives in L2 English

If the L1 pro-drophood does not make it difficult for Japanese learners of English to acquire overt subjects, we would predict that they wouldn't have difficulty acquiring English expletive constructions, either. Indeed, expletive pronouns (pleonastic *it* and *there*) are not difficult for them to acquire, except the *it* construction with *seem/ appear*, according to Yoshimura and Nakayama (2010b).

The expletives *there* and *it* appear in the subject position in English in order to meet an EPP requirement in English, which requires the sentence to have a subject in [Spec, TP] in the tensed construction.

(9) a. $[_{TP}[_{Spec}$ *There*$][_{T'}[_{VP}$ *is a man in the garden*]]].

b. $[_{TP}[_{Spec}$ *It*$][_{T'}[_{VP}$ *is said that dogs are more friendly than cats*]]].

However, there is a crucial syntax-semantic difference between the two expletives, as illustrated in (10) and (11), namely, which DP (Determiner Phrase) the verb agrees with in the sentence.

(10) a. *There exist no good solutions to this problem.*

b. **There exists no good solutions to this problem.*

(11) a. *It seems at this point equally possible that he'll resign and that he'll stay in office.*

b. **It seem at this point equally possible that he'll resign and that he'll stay in office.*

The *there* examples in (10) show that the verb agrees with the post-verbal DP, whereas the *it* examples in (11) indicate that the verb agrees with the expletive in

the subject position, not the post-verbal CP (Complementizer Phrase).[13] In contrast, Japanese does not have overt expletives similar to *there* and *it*, as illustrated below.

(12) a. *There is* [$_{\text{DP}}$ *a strange man*] *in the garden.*

b. [$_{\text{DP}}$ *Siranai otoko*] *ga* *niwa ni iru.*
unknown man NOM garden in exist

(13) a. *It is possible* [$_{\text{CP}}$ *to go to the moon*].

b. [$_{\text{CP}}$ *Tuki ni iku no*] *wa* *kanoo* *da.*
moon to go NMLZ TOP possible COP

(14) a. *It is said* [$_{\text{CP}}$ *that John is a big liar*].

b. [*Zyon* *wa* [$_{\text{CP}}$ pro *oousotuki da to*] *iwarete iru.*
TOP big liar copula COMP is said

The post-verbal DP *a strange man* in (12a) appears in the subject position as *siranai otoko* in (12b); similarly, the *to*-infinitive in (13a) and the *that*-CP in (14a) are parallel to the CPs in the subject position, headed by the nominalizer *no* in (13b) and the complementizer *to* in (14b), respectively. In these structures, *there* and *it* appear in [*Spec*, TP] in English, whereas a lexical DP and a CP appear in the subject position in Japanese. Thus, the common assumption is that Japanese does not have an overt expletive on par with *there* or *it* in English, and no overt expletive constructions in the language.[14] Because of these being no counterparts in L1, we would expect some difficulty for Japanese learners of English to acquire this type of subject. On the other hand, English-speaking learners of Japanese have little difficulty producing the equivalent Japanese sentences because they do not differ structurally from other constructions.

Of 24 expletive *there*- and 22 expletive *it* sentences identified in Yoshimura and Nakayama's (2010a) compositions, only one erroneous expletive construction (15b) was found, according to Yoshimura and Nakayama (2010b). The production data suggest that even Japanese EFL learners with lower proficiency did not have problems producing English expletive constructions.

**13** The sentences in (10) and (11) are taken from McCloskey (1991: (3) (18) (19)) (see also Safir 1985). Note that coordinate clauses agree with the verb *seem* when they appear in the subject position, as pointed out in McCloskey (1991: 564).

(i) *That he'll resign and that he'll stay in office seem at this point equally possible.*

**14** It is a controversial issue whether or not Japanese permits expletive *pro* constructions on a par with (13a) and (14a). However, there seems to be no theoretical reason to posit that this null possibility should be excluded in the language.

(15) a. **It cause lower communication*. (Low Group)

b. **There are even a cell fone which you can use abroad*. (High Group)

Yoshimura and Nakayama (2010b) also asked 20 native speakers of English and 16 Japanese EFL learners the acceptability of the sentences with expletive *there* and *it* by employing a magnitude estimation task (Bard, Robertson, and Sorace 1996).[15] For the sake of developmental comparison, the learners group was also separated based on their TOEIC scores (Lower Proficiency Group: $n$ = 8, TOEIC Ave. 575, SD 146.53; Higher Proficiency Group: $n$ = 8, TOEIC Ave. 834, SD 78.58). The average TOEIC scores by the two groups were significantly different. Out of 45 test sentences, 16 were relevant expletive sentences including (16c), which was from Kuribara (2003).

(16) a. *There were many buildings that fell due to the Great Kobe Earthquake.*

b. *It appears that our students danced all night to celebrate their graduation.*

c. **This time Ø seems that he followed my advice.*

The results of this acceptability judgment suggested that the learners could, generally speaking, discriminate the "good" and "bad" sentences like the natives, except the *it* construction with *seem/appear* (e.g., (16b)). Because the *seem/appear* counterparts are not verbal constructions in Japanese (e.g., adjectival *rasii* and adjectival noun *yoo*), Japanese EFL learners have to learn the semantic property of these verbs in English. Namely, no thematic (semantic) role can be assigned to the subject position, resulting in the occurrence of *it* (i.e., with the tensed clause complement) in this case. In addition, these verbs involve raising constructions for a Case reason if the lower CP is infinitive, as in (17) (where *t* stands for a trace).

(17) a. *She*$_i$ *seems* [$t_i$ *to be happy*].

b. *There*$_i$ *appear* [$t_i$ *to exist many millionaires in China*].

Given these syntactic-semantic distinctions coupled with the L1–L2 discrepancy, it is not strange at all even if it takes time for Japanese learners to acquire the *seem/appear* constructions. Therefore, they performed poorly in the judgment study. Despite the slow lexical learning of these raising verbs, Japanese college students

**15** The participants were asked to rate the acceptability of the sentences with respect to the norm sentence: *We walk to the station every morning*. Their raw scores were then log-converted (i.e., 1 being as acceptable as the norm sentence) and compared. The example log-converted scores by the native speakers are: *Lee's dog barked at me* (Log-converted score 1), *Well grew babies* (Log-converted score 0.3). As shown, the grammatical and the ungrammatical sentences exhibit different scores. The lower the number, the greater the unacceptability.

can acquire the knowledge (i.e., the EPP-driven constraint) for the presence of an overt subject in [Spec, TP] as well as nominative Case-checking requirement via the feature [TENSE] in T.[16] This is strong evidence to support the claim that narrow syntax is not difficult for L2 learners to acquire.

# 3 Narrow syntax and syntax-semantics/syntax-pragmatics interfaces

## 3.1 Japanese learners' acquisition of WH-questions in L2 English (Narrow syntax)

Is WH-movement difficult for Japanese learners of English to acquire because Japanese doesn't require WH-words to be at the sentence initial position when uttered?[17] Generally speaking, learners do not seem to have difficulty producing WH-questions. Yoshimura and Nakayama (2010a) observed only one error out of 25 WH-questions in the compositions by the 44 Japanese college students in study abroad contexts.

(18) *Therefore, we have to consider that how to use discoveried and inventions.* (Low Group)

However, Hawkins and Hattori (2006) claim that what learners are doing is not WH-movement, but rather WH-scrambling. They assume that in order to acquire English WH-movement, Japanese learners of English must understand the following two properties of Move α in the target language.[18]

(19) a. One WH-word/phrase must appear in the matrix [Spec, CP];

b. Such movement must observe the Attract Closest Principle.[19]
*Attract Closest Principle* (Radford 2004: 162)
A head which attracts a given kind of constituent attracts the closest constituent of the relevant kind.

---

**16** In addition, the students can understand that the associated DP, being semantically the subject, must be moved to adjoin to the expletive *there* in [Spec, TP] as *there* is a LF affix (Chomsky 1991; McCloskey 1991; Lasnik 1995). This means that covert LF movement is not difficult for Japanese EFL learners to acquire in this case.

**17** For processing of English WH-interrogatives by Japanese EFL learners and Japanese WH-interrogatives by English-speaking JFL learners, see Aoshima, Phillips and Weinberg (2004) and Lieberman, Aoshima and Phillips (2006), respectively.

**18** See Hawkins and Hattori (2006: 280 (16)) for their exact stipulation of the requirements for WH-movement. See also Miyamoto and Okada (2004) and Umeda (2005).

**19** The Attract Closest Principle is proposed by Radford (2004: 162) in order to account for Chomsky's (1973) superiority effects as in (21b). Hawkins and Hattori (2006) assume, following Radford (2004, 2009), that this principle is a constraint on syntactic movement.

For example, the contrasts in grammaticality in (20) and (21) can be accounted for by (19a) and (19b), respectively.

(20) a. *What*$_i$ *do you think* $t_i$ *John bought* $t_i$ *yesterday?*

b. **Do you think John bought what yesterday?*

(21) a. *What*$_i$ *did you say* $t_i$ *the students ate* $t_i$ *where?*

b. **Where*$_i$ *did you say* $t_i$ *the students ate what* $t_i$*?*

(20b) induces a violation of (19a) because *what* remains in-situ, and (21b) results in a violation of (19b) because *what* is closer to the matrix [Spec, CP] than *where*. In contrast, the scrambling of WH-phrases in Japanese is immune to such constraints.

(22) a.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (*Anata wa*) | [*Zyon ga* | *kinoo* | *nani o* | *katta* | *to*] | *omoimasu* | *ka.* |
| you TOP | John NOM | yesterday | what ACC | bought | COMP | think | Q |

‘What do you think that John bought yesterday?’

b.

| | | | | | | |
|---|---|---|---|---|---|---|
| *Doko de* | (*anata wa*) | [*gakuseitati ga* | *nani o* | *tabeta to*] | *iimasita* | *ka.* |
| where | you TOP | students NOM | what ACC | ate COMP | said | Q |

‘What did you say the students ate where?’

As shown in the translation, although the sentences in (22) correspond to (20b) and (21b), respectively, they are grammatical with *nani* ‘what’ remaining in the object position of the embedded clause and *doko de* ‘where’ being moved in the sentence initial position, respectively. The single clause WH-interrogative like *what did you buy at the store?* would have the representation of (23b), not the structure of (23a), if we were to follow the WH-scrambling hypothesis.

(23) a. [$_{CP}$ [$_{SPEC}$*What*$_i$] [$_{C'}$ *did* [$_{IP}$ *you buy* <*what*$_i$> *at the store*]]]*?*

b. [$_{CP}$[$_{C'}$[$_{IP}$[$_{SPEC}$*What*$_i$] *did* [$_{IP}$ *you buy* $t_i$ *at the store*]]]]*?*

A crucial difference between the two is that *what* moved into [Spec, CP] in (23a), while it moved into [Spec, IP] in (23b).

Yoshimura and Nakayama (2011), on the other hand, claim that Japanese-speaking learners of English can acquire WH-movement based on their acceptability judgment experiment. The learners were able to exclude WH-in-situ correctly as in (24) (i.e., non-echo-questions), and move the WH-word into the embedded [Spec, CP] in biclausal WH-interrogative *ask*-sentences as in (25) (see also Yoshimura and Nakayama 2009, 2010a, b; Kaneko 2005; Yamashita 2007). No proficiency difference was observed in (25a) vs. (25b). However, they were unable to move a WH-word

in the matrix [Spec, CP] in biclausal WH-interrogative *think*-sentences as in (26) (Kaneko 2005; Radford and Yokota 2006; Wakabayashi and Okawara 2003). There was a proficiency difference observed in (26a) vs. (26b). Short-distance WH-movement posed no problem, but long-distance WH-movement was difficult for the Japanese EFL learners.

(24) a. **Susan bought <u>what</u> at the computer store?*

b. **You are saying the child broke <u>what</u> in the classroom?*

(25) a. *Did you ask <u>what</u> John was doing at the library?*

b. **<u>Which book</u> did you ask Lauren liked most?*

(26) a. *<u>What</u> do you think Jennifer bought for her mother at the store?*

b. **Did Joe think <u>who</u> prepared the dinner?*

Once the Japanese learners of English come to understand the mechanism of WH-movement, i.e. obligatorily valuing the [*u*WH] feature in [Spec, CP], the early acquisition of short-distance WH-movement follows from locality (which we assume is the core notion of UG).[20] However, learning the selectional restriction on an embedded CP (Chomsky 1973) takes time because it is an English specific rule, and in turn, it delays the acquisition of successive cyclic long-distance WH-movement.[21]

**20** The *u* in [*u*WH] means "unvalued", which needs to be valued by means of merge operation (Radford 2009).

**21** This is not compatible with Hawkins and Hattori's (2006) claim for obligatory WH-scrambling in the Japanese ESL learners' interlanguage grammars. If it were that WH-scrambling resulted from L1 transfer, such early short-distance vs. delayed long-distance distinction would not appear because Japanese permits both types of movement in its syntax. We also question how Japanese speakers can arrive at the notion of obligatoriness of WH-movement because scrambling is optional. Thus, we conclude that the [*u*WH] feature is acquirable, and Japanese learners of English can attain the representation of WH-movement in their interlanguage grammars. Then, how do we account for Hawkins and Hattori's main argument in support of the WH-scrambling analysis, i.e., Japanese learners' violation of Attract Closest Principle? Restating Pesetsky's (1987) Nested Dependency Condition as an LF filter on scope interpretations, we proposed a syntax-LF interface account for Japanese learners' difficulty with superiority effects in (21b). Our basic analysis is that movement is a syntactic operation, and UG properties are given in narrow syntax, while an LF filter is something that L2 learners need to learn based on linguistic input, with either direct or indirect evidence. In addition, computational limitation, e.g., working memory capacity in L2, may have brought the results they obtained.

## 3.2 Binding issues

When L2 acquisition studies are considered from a modularity perspective, one finds few interface studies related to LF such as syntax-semantics and syntax-discourse/ pragmatics. On the assumption that the acquisition of these interface matters is based on language use and linguistic experience, they would be more difficult to acquire than those in narrow syntax and morphophonology. Below, we first look at a bound variable interpretation, which is structurally determined, and then, discuss short and long distance binding which lies in both narrow syntax and syntax-pragmatics domains.

### 3.2.1 Bound variable interpretations

The bound variable interpretations are obtained when the antecedents of pronouns and reflexives are quantifiers. Consider the following sentences.

(27) a. *Everyone$_i$ used his$_i$ umbrella.*

b.

| | | | | | |
|---|---|---|---|---|---|
| *Daremo$_i$* | *ga* | *zibun$_i$ no* | *kasa* | *o* | *tukatta.* |
| everyone | NOM | self GEN | umbrella | ACC | used |

c. **Daremo$_i$ ga kare$_i$ no kasa o tukatta.*
his GEN

d. *Daremo$_i$ ga pro$_i$ kasa o tukatta.*
his

*Zibun* is a reflexive counterpart to *himself/herself* and *kare/kanozyo* and *pro* are called overt and null pronouns here.[22] The quantifier binds the reflexive and the pronouns because it c-commands them and they are co-indexed.

(27a) and (27b) are true in the following situation.

(28) *Larry, Bill, Brian and Robert worked for the same accounting firm. One day, after a meeting, they decided to have lunch together at a nearby restaurant. Because both Brain and Robert needed to contact their clients, they asked Larry and Bill to go to the restaurant first. Since it was raining, Larry, using his umbrella, walked to the restaurant. Bill opened his own umbrella and followed Larry. After contacting their clients, Brian walked to the restaurant using his blue umbrella while Robert went there with his compact umbrella.*

22 Although we call *kare* an overt pronoun here, it is actually a demonstrative. See Hoji (1991), Noguchi (1997), and the *Handbook of Japanese Historical Linguistics* in this HJLL series.

Larry, Bill, Brian, and Robert each used his own umbrella. This situation can be described by (27a), where the pronoun *his* refers to the individuals (the distributive reading). In Japanese, on the other hand, this situation can be described either by (27b) with *zibun* or possibly (27d) with a null pronoun (*pro*). However, so-called overt pronoun *kare* in (27c) cannot be used for the bound variable interpretation (Saito and Hoji 1983; Hoji 1991). Keeping this in mind, let us discuss the acquisition of L2 English and L2 Japanese reflexives and pronouns with the bound variable reading.

### 3.2.2 Pronouns and reflexives in L2 English

The acquisition of reflexives and pronouns has been investigated extensively among Japanese EFL learners (see below). However, the number of studies on L2 learners' bound variable readings is rather limited. For instance, Ito (2003) investigated interpretations of pronouns in a variety of syntactic structures among Japanese high school EFL learners in Japan.[23] She employed a pictorial truth value judgement task (cf. Crain and McKee 1986; Chien and Wexler 1990) and found that they allowed bound variable readings of pronouns at a rate of 80% in the test sentences like (29).

(29) *Every boy$_i$ dreamed that the man shot him$_i$.*

The learners did not seem to have much difficulty interpreting *him* with *every boy*. Oya (2006) also investigates the acquisition of the bound variable reading by Japanese high school (11th grade), and college sophomore and junior EFL learners. Her two questionnaire experiments employed a truth value judgment task with narratives. For instance, EFL learners were asked to judge whether sentences like (27a) matched stories like (28). Learners were to indicate TRUE if the test sentence matched the story, or FALSE if it did not. Since the results of two experiments were similar, only Experiment II is discussed here. Table 3 shows the correct response rates of sentences like (30a) and (30b). The number in the parenthesis indicates the number of learners. There is an increase in the correct response rate in reflexives and pronouns by learner group, but the bound variable reading was observed even among high school learners.

**23** Note that one Japanese English textbook for the 7th graders (the first year students of a junior high school) contained 59 *he/she* and 14 *him/her*, but no *himself/herself*, according to Shirahata (2007). In the 8th grader's textbook, there were 61 *he/she* and 12 *him/her*, and one *himself/herself* and in the 9th grader's textbook, they were 72 *he/she*, 23 *him/her*, and one instance of *himself/herself*.

**Table 3:** Correct response rates of the bound variable reading in Oya's Experiment II

| | High Sch | Col. Soph | Col. Junior |
|---|---|---|---|
| (30a) *Everyone praised himself.* | 77%(21) | 93%(16) | 94%(14) |
| (30b) *Everyone washed her spoon.* | 79%(18) | 82%(10) | 92%(12) |

Similar to Ito's study, Oya's high school students scored close to 80% on the bound variable reading and the correct response rate improved with study. The bound variable reading does not seem difficult for Japanese EFL learners to acquire. This may be positive L1 transfer because this reading exists in Japanese. Once the structure is acquired and pronouns and reflexives are identified, this reading becomes available as it is determined structurally.

Since the distributive readings may be easily obtained in Ito's and Oya's studies due to the "every-his/her" agreement in number, Nakayama, Wakabayashi, and Hosoi (2006) investigated the bound variable reading with sentences with *all* like (31) by using a questionnaire with a truth value judgment task.

(31) *All children washed their spoons.*

They found that Japanese college EFL learners correctly took bound variable readings 97% of the time. These studies indicate that the learners acquired the bound variable reading quite early because English has only overt pronouns, i.e., no negative transfer, even though Japanese *kare* and *karera* cannot have BV readings. It is predicted that the reverse should not be so simple since English-speaking learners of Japanese must learn the properties of the overt and null pronouns as well as *zibun* in Japanese.

### 3.2.3 *Zibun* in L2 Japanese

The complexity of anaphoric expressions and the differences found in English and Japanese made the study of L2 grammar acquisition stimulating, especially, from the perspective of transfer. Consider the sentences in (27) with (32) below.

(32) a. *All (people)$_i$ used their$_i$ (own) umbrellas.*

b.

| | | | | | | |
|---|---|---|---|---|---|---|
| *Minna$_i$* | *ga* | *zibun$_i$* | *no* | *kasa* | *o* | *tukatta.* |
| all people | NOM | self | GEN | umbrella | ACC | used |

c. **Minna$_i$ ga* *kare$_i$ no kasa o tukatta.*
his (intended meaning: 'Everyone used his own umbrella.')

d. *Minna$_i$ ga* *pro$_i$ kasa o tukatta.*
their

As in (27b), *zibun* can appear in the possessive position where English reflexives cannot. When *his* refers to *everyone* in (30b), it can evoke the relevant bound variable interpretation. As we saw in (28), Larry, Bill, Brian, and Robert each used his own umbrella. This situation can be also described by (32a), where the pronoun *their* refers to the individuals. In Japanese, on the other hand, this situation can be described either by (27b) or (32b) with *zibun*. However, *kare* in (27c) and (32c) cannot be used for the bound variable interpretation while *pro* can as in (32d).

Kano and Nakayama (2004b) and Nakayama and Kano (2007) employed the truth value judgment task with narratives to investigate whether English-speaking JFL learners could interpret *zibun* and *zibuntati* as a bound variable. Stories like (28) with test sentences like (32b) were used on the bound variable reading. In Kano and Nakayama (2004b), JFL learners were divided into three groups, ACTFL Intermediate Mid (n = 18), Intermediate High (n = 5), and Advanced (n = 7), and all levels achieved a very high accuracy rate for accepting the bound variable readings (97%, 100%, and 100%, respectively; cf. Native Speakers of Japanese 96%).[24] This along with the findings from the interpretation of *zibuntati* in Nakayama and Kano (2007) provided strong evidence for the idea that the learners take the bound variable interpretation as the null hypothesis and apply it whenever a binding configuration holds. If this scenario is correct, JFL learners might interpret *kare/kanozyo* as a bound variable, although this is equivalent to what L1 transfer predicts.

### 3.2.4 Pronouns in L2 Japanese

As seen above, Japanese permits overt and null pronouns. Studies such as Kanno (1997, 1998) seem to show evidence for English-speaking JFL learners' native-like ability in the interpretation of overt and null pronouns from an early stage of learning. However, Masumoto (2008) and Pimentel and Nakayama (2012a, b) present counterevidence to Kanno's results and show that English-speaking JFL learners with lower proficiency accept overt pronouns with the bound variable reading.

Kanno (1997) investigates English-speaking JFL learners' knowledge of the contrast between null and overt pronouns. An example sentence with *kare* 'he' and *dare* 'who' is listed below.

(33) *Dare$_i$ ga* [*kyoo kare$_i$ ga* *osoku naru*] *to* *itte iru n desu ka?*
who NOM today he NOM late become that is saying COP Q
'Who is saying that he would be late today?"

Q: *Dare$_i$ ga* *kyoo* *osoku naru n* *desyoo* *ka?*
who NOM today late become probably Q
'Who do you suppose will be late today?"
(a) same as *dare* (b) <u>another person</u>

24 Note that the Intermediate Mid level is the level at which *zibun* is first introduced.

The participants were instructed to indicate whether the subject argument in the embedded clause referred to (a) the same as *dare* or (b) another person. They were also instructed that they had a third option of choosing both (a) and (b) if they thought this was appropriate. The following table illustrates the results of the interpretations for both JFL learners (JFL), and native speakers (NS) for bound variable readings (i.e., (a) and (a) & (b) for the Null and Overt pronoun (Q) questions) as well as referential readings (R).

**Table 4:** (a) and (a) & (b) answers in Kanno (1997)

| | Null pronoun (Q) | Overt pronoun (Q) | Null pronoun (R) | Overt pronoun (R) |
|---|---|---|---|---|
| JFL (*n* = 28) | 78.5% | 13% | 81.5% | 42% |
| NS (*n* = 20) | 83% | 2% | 100% | 47% |

Kanno's results show that the percentages of responses by the JFL and the NS groups that were exclusively (b) in the test sentences were 21.5% and 17%, respectively. The difference between the two groups was not statistically significant. In the sentences containing overt pronouns with quantified noun phrases, 98% of the NS group's responses and 87% of the JFL group's responses were answer (b) only. From these results Kanno concluded that the JFL learners had knowledge of the Overt Pronoun Constraint (Montalbetti 1984: Overt pronouns cannot link to formal variables iff the alternation overt/empty obtains), which is assumed to be in UG, and therefore, she concluded that they had access to UG (cf. Bley-Vroman 1989; Clahsen and Muysken 1989).

In a follow-up study Kanno (1998) tested twice, at the beginning of the semester and 12 weeks later, whether participants accepted an interpretation of the overt pronoun *kare* that referred to (a) the subject antecedent, (b) a sentence-external antecedent, or both (a) and (b). The JFL learners more readily chose a quantifier antecedent for null pronouns over overt pronouns with a referential antecedent. The results were similar to those of Kanno (1997) (see Table 5 below). Thus, she concluded that the learners had correct knowledge of overt and null pronouns in Japanese.[25]

Masumoto (2008) and Pimentel and Nakayama (2012a), however, show different results. They examined American English-speaking JFL learners with a questionnaire with the truth value judgment task. For instance, Pimentel and Nakayama (2012a) examined the following sentence, and their results are also given in Table 5.

**25** Sheen (2000) disagrees with Kanno (1998) because of the increase in the incorrect bound variable interpretations from Session 1 to Session 2 (29% to 34%).

(34) *Dono itoko* *mo* *kare no* *imooto* *o* *yonda.* (overt pronoun)
which cousin also he GEN younger sister ACC called
'Every cousin called his younger sister.'

Because the tasks employed in Kanno (1997, 1998), and Masumoto (2008), and Pimentel and Nakayama (2012a) were different, Pimentel and Nakayama (2012b) employed Kanno's methodology with her and their test sentences. The results summarized in Table 5 include only comparable JFL learners. The percentages indicate the Overt Pronoun Constraint violations, i.e., overt pronouns referring to the quantifier antecedents, by learners with comparable proficiency.

**Table 5:** Percentage of Overt Pronoun Constraint violations

| | Kanno | | | Masumoto | Pimentel and Nakayama | |
|---|---|---|---|---|---|---|
| | (1997) | (1998) | | (2008) | (2012a) | (2012b) |
| | | Beginning | 12thwk | | | |
| Intermediate Low | 13% | 29% | 34% | 61% | 56% | 58% |

As shown, Kanno's (1997) learner group had the lowest number of the Overt Pronoun Constraint violations (13%). In comparison, Pimentel and Nakayama (2012a, b) show Overt Pronoun Constraint violations of 56% and 58%, respectively, while the L2 learners in Masumoto (2008) showed a violation of 61%. This suggests that the learners at this level do not grasp the fact that overt pronouns cannot have the bound variable reading until a more advanced stage of learning, though the correct referential interpretations were available from early stages of learning. Masumoto (2008) and Pimentel and Nakayama (2012a, b) suggest L1 transfer (or the bound variable reading is automatically obtained once the structure is acquired) as a reason. Why do these studies differ from Kanno's? Pimentel and Nakayama (2012b) point out that learners' exposure to Japanese outside the classroom may have affected their knowledge (Hawaii vs. Ohio). Pimentel and Nakayama's learners improved their understanding in the following level (Intermediate Mid) where *zibun* was introduced. The learners must have adjusted their understanding of referentially dependent expressions (overt and null pronouns and *zibun*) in Japanese (cf. Feature Assembly Hypothesis of Lardiere 2005). In other words, Kanno's learners could have actually been more advanced in this respect than Masumoto's and Pimentel and Nakayama's learners. If this is correct, Kanno's learners might have had L1 transfer or the default bound variable reading with an overt pronoun before the stage Kanno examined. If so, we could conclude that learners with low proficiency treat the

quantifiers as viable antecedents for the overt pronouns (cf. Pérez-Leroux and Glass 1999).[26]

In sum, these studies suggest the early bound variable readings in L2 English and Japanese, and the learners learn that *kare* cannot take the quantifiers as their antecedents. Since the bound variable reading is structurally defined, the reading becomes available once a noun is labeled as a referentially dependent pronominal. Learning the lexical properties of *kare* is not simple because pragmatics and discourse are involved. Thus, the lexical learning of *kare* takes time.

## 3.3 Referential antecedents and binding domain

In the last more than a couple decades, many L2 acquisition studies on binding have been conducted within the Principles-and-Parameters framework. In particular, considerable attention has been paid to the issue of whether L2 learners can reset a parametric value relevant to the binding domain (governing category) in interpreting the anaphoric relationship between a reflexive pronoun and its antecedent. A basic assumption has been put forward that L2 learners often encounter difficulty in interpreting the anaphoric relationship between a reflexive pronoun and its antecedent due to their failure in appropriate parameter resetting, i.e., L1 transfer (e.g., Finer and Broselow 1986; Hirakawa 1990).

However, do L2 learners indeed reset their parameter during the course of L2 acquisition? In this section, we summarize L2 acquisition studies on English reflexives by L1 Japanese, Korean, and Chinese-speaking learners, and then the acquisition of Japanese *zibun* by L1 English, Chinese and Turkish-speaking learners. In particular, we discuss why a parametric approach to the binding domain (governing category), as proposed in Manzini and Wexler (1987) and Wexler and Manzini (1987), does not work.

### 3.3.1 *Himself/herself*, *zibun*, *caki*, *ziji*, and *kendi*

It has been well documented that Japanese *zibun*, Korean *caki*, Chinese *ziji* and Turkish *kendi* differ from English *himself* in that unlike the latter, the former group of reflexives permits long-distance binding as well as short-distance binding. Observe (35) and, in contrast, (36).

---

**26** Note that the fact that Masumoto/Pimentel and Nakayama studies did not support Kanno's results doesn't necessarily mean that the UG availability Kanno supported is rejected. Since Japanese overt pronouns are actually demonstratives, whether or not our discussion rejects the existence of the Overt Pronoun Constraint in UG is not relevant here, either.

(35) *John$_i$ thought Tom$_j$ blamed himself$_{*i/j}$.*

(36) a.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Taroo$_i$* | *ga* | *Kazu$_j$* | *ga* | *zibun$_{i/j}$* | *o* | *semeta* | *to* | *itta.* |
| Taro | NOM | Kazu | NOM | self | ACC | blamed | COMP | said |

‘Taro said that Kazu blamed self.’

b.

| | | | | |
|---|---|---|---|---|
| *John$_i$-i* | *Mary$_j$-i* | *caki$_{i/j}$-lulu* | *coahan-n-ta-ko* | *malha-yess-ta.* |
| John NOM | Mary NOM | self ACC | likes COMP | said |

‘John$_i$ said that Mary$_j$ likes himself$_i$/herself$_j$.’

c.

| | | |
|---|---|---|
| *Zhangsan$_i$* | *renwei* | *Lisi$_j$ xiangxin ziji$_{i/j}$.* |
| Zhangsan | think | Lisi trust self |

‘Zhangsan thinks Lisi trusts self.’

d.

| | | | | | |
|---|---|---|---|---|---|
| *Ali$_i$-* | *Veli$_j$-* | *kendi-si$_{i/j}$-ni* | *sucla-di* | *diye* | *dusun-du.* |
| Ali-NOM | Veli-NOM | self-3sg-ACC | criticize-PST | COMP | think-PST |

‘Ali thought that Veli blamed himself/him.’

In (35), the reflexive *himself* must take the embedded subject *Tom* as its antecedent, and cannot take the matrix subject *John*. The English reflexive permits only short-distance binding, not long-distance binding. As shown in (36), on the contrary, the Japanese reflexive *zibun*, the Korean *caki*, the Chinese *ziji*, and the Turkish *kendi* can take the matrix subject as well as the embedded subject as its antecedent.

Reflexive pronouns are subject to Principle A (An anaphor is bound in its governing category) of the binding theory (Chomsky 1981). This means that an anaphor must be c-commanded by and co-indexed with an antecedent in its governing category. As (35) and (36) illustrate, however, governing category seems to vary from language to language. In dealing with this “parametric variation”, Manzini and Wexler (1987: 419) propose the Governing Category Parameter, as stated in (37):

(37) α is a governing category for β iff α is the minimal category which contains β and has
a. a subject; or
b. an Infl; or
c. a Tense; or
d. a ‘referential’ Tense; or
e. a ‘root’ Tense.

The principle subsumes the view that languages differ in the size of a governing category in which reflexives are bound: The (37a) setting is for *himself* as in (35), the most restricted or local domain, while the (37e) setting is for *zibun*, *caki*, *ziji* and *kendi* as in (36), the least restricted or non-local domain. As such, this notion is crucial for the parameter resetting approach to L2 acquisition of short-distance vs.

long-distance reflexive binding (see also Wexler and Manzini 1987). That is to say, English-speaking learners of Japanese must reset their governing category domain value (37a) to (37e) in their L2 Japanese grammar while Japanese-speaking learners of English must reset the value (37e) to (37a) in their L2 English grammar. If there is L1 transfer, English-speaking learners of Japanese permits short-distance binding, but not long-distance binding initially, whereas Japanese-speaking learners of English allows both short-distance and long-distance binding. The former is the undergeneration case while the latter is the overgeneration case.

### 3.3.2 L2 English *himself/herself*

Table 6 is a summary of the results of three previous studies relevant to the present discussion of L2 English reflexive acquisition (where SD and LD stand for short-distance and long-distance, respectively).

**Table 6:** Interpretations of English reflexives by Japanese and Korean learners

| | Participants | English proficiency | L1 | Tensed Clause | | | Infinitive Clause | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | SD | LD | Either | SD | LD | Either |
| Finer & Broselow (1986) | Adults (n = 6) | – | K | 91% | 8.3% | 0 | 58% | 37% | 4.2% |
| Hirakawa (1990) | Adults (n = 65) | – | J | 77% | 17% | 5.9% | 55% | 36% | 7.8% |
| Thomas (1991) | Adults (n = 70) | Low (n = 20) | J | 80% | 5% | 5% | – | – | – |
| | | Mid (n = 25) | J | 76% | 0 | 16% | – | – | – |
| | | High (n = 25) | J | 84% | 0 | 16% | – | – | – |

Finer and Broselow (1986) find that Korean adult learners chose an intermediate parametric value between L1 Korean and L2 English in their interpretation of English reflexives. Similarly, Hirakawa (1989, 1990) claim that Japanese high school and college students accepted non-local antecedents for the English reflexive. Thomas (1991) suggested that Japanese adult learners of L2 English seemed able to reset the relevant parameter, hence there was no serious L1 transfer. Of relevance to the present discussion is that these results pointed to the two general acquisition patterns of reflexive binding: ESL learners can interpret short-distance binding of English reflexives far more correctly than long-distance binding, and second, they can perform much better in tensed clauses than in infinitive clauses. Put simply, the results suggest that parameter setting is not a key principle underlying the L2 acquisition of reflexive binding.[27]

27 Wakabayashi (1996) and Watanabe et al. (2008) considered parametric values in (37) among Japanese ESL learners' interlanguage grammars. For instance, the latter study found 27 learners with the English value, 7 learners with the Japanese value, and 48 learners with the Russian value, which supports the UG sanctioned values at the interlanguage stages. See Thomas (2006) for an overview.

Employing the truth value judgment task, Yoshimura et al. (2012) examine whether L2 learners indeed reset their parameter during the course of acquiring L2 reflexive binding. We formulated three predictions:

(38) a. If L2 learners start with their L1 parametric value on the parameter resetting approach, L2 learners of English with low proficiency would permit both short-distance and long-distance binding more often than those with high proficiency.

b. If locality is a core notion of anaphor binding, short-distance binding would not pose any problem for L2 English learners, regardless of their L1s.

c. If L2 learners do not have sufficient syntactic knowledge in English, long-distance binding would be delayed in L2 acquisition due to L1 discourse-pragmatic transfer.

**Table 7:** Correct interpretations of English reflexives

| | | | Tensed Clause | | Infinitive Clause | |
|---|---|---|---|---|---|---|
| Participants | English proficiency | L1 | SD | LD | SD | LD |
| Adults (n = 205) | Low (HS1) (n = 15) | J | 93% | 47% | 93% | 62% |
| | Mid (HS2) (n = 25) | J | 96% | 75% | 96% | 71% |
| | Mid (JC1) (n = 62) | J | 84% | 69% | 94% | 66% |
| | High (T) (n = 25) | J | 93% | 88% | 88% | 90% |
| | Low (KC1) (n = 28) | K | 95% | 65% | 95% | 52% |
| | Mid (KC2) (n = 24) | K | 85% | 79% | 96% | 67% |
| | High (CC) (n = 26) | C | 90% | 64% | 98% | 86% |

Table 7 shows the mean percentages of short-distance and long-distance correct responses from Japanese EFL learners (HS1 = high school first year, HS2 = second year, JC = college first year, T = English teachers), Korean EFL learners (KC1, KC2), and Chinese ESL learners at a Canadian university (CC). All the learner groups accepted short-distance as successfully as native speakers of English, although short-distance was easier in infinitive clauses than in tensed clauses. On the other hand, long-distance showed no significant differences on clause types, but revealed a significant main effect on groups and interaction. Moreover, long-distance was difficult for the HS1, JC, KC1, and CC groups in tensed clauses whereas it was difficult for HS1, JC, KC1, and KC2 groups in infinitive clauses. These results suggest that long-distance is more difficult to acquire than short-distance. In addition, the results

suggest that in order to understand English reflexive binding in the infinitive clause L2 learners must be more proficient than the intermediate (Mid) level.[28]

According to the results, prediction (38a) is not supported while prediction (38b) is borne out. Therefore the parameter resetting approach cannot be a plausible proposal for L2 reflexive acquisition. Instead, we proposed an alternative account for the short-distance – long-distance asymmetry based on the well-accepted view that *zibun*, *caki*, and *ziji* can be either anaphoric or logophoric (e.g., Kuno 1973; Sells 1987; Abe 1997, among others). The assumption is that an anaphoric reflexive must be syntactically construed as being coreferential to its antecedent while a logophoric reflexive refers to a person whose thought, state of consciousness, or point of view is being reported (Clements 1975; Stuart 2003; Kuroda 1973 for *zibun*). More specifically, by adopting a syntax-pragmatics interface approach (Sorace 2007) together with the modular structure of grammar (Chomsky 1995), a short-distance reflexive is subject to Binding A in narrow syntax, but a long-distance reflexive is subject to a relevant pragmatic principle in discourse-pragmatics.[29]

Given this crucial distinction, we can explain why the rejection of long-distance binding is delayed in L2 English. These L2 learners need to figure out whether their prior linguistic experience should or should not apply to English reflexives at hand. Our speculation is that they need sufficient input before arriving at a firm understanding of the irrelevance of the L1 pragmatic principle in L2, hence presumably the delayed rejection of long-distance reflexive binding. Note importantly that this does not necessarily entail L1 transfer. Since L2 learners whose L1 is Japanese, Korean, or Chinese know the existence of long-distance reflexive binding through their prior language learning, they take time in reaching a decision about the relevance or the irrelevance of the given pragmatic principle to the new language in the course of L2 acquisition. In short, they learn from trial and error, and it is not necessarily the case that they apply their L1 pragmatic knowledge.[30]

---

**28** The CC group's performance was poorer than that of the Control group in the tensed condition while it was as good as that of the Control group in the infinitive condition. However, the results may have been affected by four participants who rather consistently gave incorrect responses, because their removal brought the difference between the CC and the Control groups not significantly different on both tensed and infinitive conditions.

**29** Due to the space limitation we do not provide a discussion of the availability of a similar logophoric interpretation of the Korean LD *caki* (Kim 1992) and the Chinese *ziji* (Huang and Liu 2001). See Yoshimura et al. (2012).

**30** The tensed vs. non-tensed asymmetry emerged due to L2 learners' insufficient syntactic knowledge. The biclausal structure (iia) was incorrectly analyzed as a monoclausal structure (iib) with the second DP being an indirect object of the main verb. This monoclausal analysis seems to have forced the matrix subject DP to mistakenly function as the antecedent of *himself*.

(i) **Dave$_i$ advised Ralph to talk to himself$_i$.*

(ii) a. *[DP$_i$ [$_{VP}$ advised DP$_j$ [PRO$_j$ to talk to himself$_j$]]]*
b. *[DP$_i$ [$_{VP}$ advised [$_{V'}$ DP$_j$ to talk to himself$_i$]]]*

The results of this study indicate that this misanalysis tended to be gradually overcome as the learners' proficiency improved.

The results revealed two significant asymmetries, early short-distance vs. delayed long-distance and early tensed vs. delayed non-tensed clauses, in the acquisition of reflexive binding. Now, let us look at L2 Japanese in the following section, which further supports our position.

### 3.3.3 L2 Japanese *zibun*

Thomas (1991) discusses the acquisition of short-distance (SD) and long-distance (LD) binding of *zibun* by English-speaking and Chinese-speaking learners of Japanese. The following table summarizes her results.[31]

**Table 8:** Correct *zibun* interpretations by 41 Japanese learners in Thomas (1991)

| L1 | Japanese Proficiency | Short-distance | Long-distance | Either |
|---|---|---|---|---|
| English | Low | 37.5% | 12.5% | 12.5% |
| | Mid | 83.3% | 0.0 | 8.3% |
| | High | 23.1% | 7.7% | 30.8% |
| Chinese | | 25% | 50% | 0% |

Thomas (1991) interpreted an increase in the either short-distance or long-distance interpretation of *zibun* to 30% at the high proficiency level of L1 English learners as indicating that the more exposed to Japanese they were, the better understanding they gained of long-distance binding. However, she left open the question of whether interpretation of *zibun* by L1 Chinese speakers is due to their preference or parameter setting, given the small size of the group ($n$ = 8). In short, these results seem to suggest a *himself-zibun* difference from the subset to superset direction (the Subset Principle) in L2 acquisition.

Shirahata (2002) conducted a longitudinal L2 study of the interpretation of *zibun* by 12 L1 English-speaking children living in Japan. Their arrival ages in Japan ranged from five to nine. The data were elicited from interviews individually administered to the children once every 2–4 months. The research designs were the same as those in Shirahata and Ishigaki (2001) and the stimulus sentence like (39) was associated with the picture and a relevant short story.

(39) *Kuma wa neko ga zibun no omotya o kowasite iru no o miteru kana.*
bear TOP cat NOM self GEN toy ACC break is NMLZ ACC looking is Q
'Is the bear looking at the cat breaking self's toy?'

**31** See also Thomas (1995) that tested *zibun*'s LF-movement analysis with a truth value judgment task.

Nine children acquired short-distance binding earlier than long-distance binding, whereas three children acquired short-distance and long-distance binding around the same time. The results were consistent with those of L1 children in Shirahata and Ishigaki (2001), confirming that overall, short-distance binding of *zibun* is acquired earlier than long-distance binding, irrespective of L1 or L2.[32]

Yoshimura et al. (2012) investigated why short-distance *zibun* binding is earlier than long-distance *zibun* binding among L2 learners, and whether parameter resetting is an appropriate approach to the acquisition of *zibun*. We employed the truth-value-judgment task and examined Chinese and English-speaking learners. Sample test situations and sentences are stated in (40). (40a) is for short-distance *zibun* binding and (40b) is for long-distance *zibun* binding.

(40) a. Narrative for short-distance *zibun* binding
*Taro and Yasuo are twin brothers. They started making plastic models one week ago: Taro's been making a model ship while Yasuo a model airplane. They finally finished making them yesterday. But, today Yasuo's model plane was broken and the broken plane was found in a trash can.*
*Taro: Why did you throw your model plane away?*
*Yasuo: I threw it away because I don't need it any longer.*
*Taro: Well, you'll be scolded by Mom.*
Test sentence (TRUE)

| | | | | | |
|---|---|---|---|---|---|
| *Taroo* | *wa* | *Yasuo*$_i$ | *ga* | *zibun*$_i$ *no* | *puramoderu o* |
| Taro | TOP | Yasuo | NOM | self GEN | plastic model-ACC |

| | | | |
|---|---|---|---|
| *gomibako ni* | *suteta* | *to* | *sirimasita.* |
| trash box into | threw away | COMP | found out |

'Taro found out that Yasuo$_i$ had thrown away his$_i$ plastic model into the trash box.'

---

**32** In addition, Kano and Nakayama (2004a) investigated the interpretation of *zibun* in (i), among English-speaking learners of Japanese and found that overall the learners tended to accept the short-distance subject as a viable antecedent candidate (87%) while the long-distance subject as an antecedent of *zibun* (27%).

(i) *Suzuki-san ga Tanaka-san ga zibun no konpyuutaa o tukatta koto o hanasita.*
Mr.-NOM Mr.-NOM self GEN computer ACC used COMP ACC said
'Mr. Suzuki said that Mr. Tanaka used self's computer.'

Kano and Nakayama also included other sentence types with *zibun de* 'by oneself' and the empathy-loaded verb *kureru* 'give'. Lower proficiency groups also failed to detect the empathy constraint on *zibun* in the empathy-locus position of the verb *kureru*.

b. Narrative for long-distance *zibun* binding
*Hanako is planning Prof. Smith's birthday party next Saturday. She e-mailed Ken on that matter this morning, but there was no reply. So she telephoned him.*
*Hanako: Did you read my e-mail?*
*Ken: Sorry, I was busy and have not checked my e-mails*
*Hanako: I see. I plan to have Prof. Smith's birthday party next Saturday, but can you come?*
*Ken: I think so, but what time?*
*Hanako: 5 pm. It's 3000 yen. Since it's a surprise party, please don't tell Prof. Smith.*
Test sentence (TRUE)

| | | | | | |
|---|---|---|---|---|---|
| *Hanako$_i$ wa Ken ga* | *zibun$_i$ no* | *meeru o* | *yonda* | *ka* | *tazunemashita.* |
| Hanako TOP Ken NOM | self GEN | mail ACC | read | Q | asked |

'Hanako$_i$ asked if Ken had read her$_i$ email.'

Table 9 shows the mean percentages of L2 learners' correct short-distance and long-distance "True" responses to the matched cases and their correct "False" responses to the mismatched cases in each L1 group.

**Table 9:** Mean percentages of correct short- (SD) and long-distance (LD) binding interpretations

| | | | Short-distance | | Long-distance | |
|---|---|---|---|---|---|---|
| L1 | Participant # | Proficiency | TRUE | FALSE | TRUE | FALSE |
| Chinese Y1 | n = 15 | IntLow | 86.7% | 93.3% | 55.6% | 75.6% |
| Chinese Y2 | n = 19 | IntMid | 96.5% | 93.0% | 78.9% | 93.0% |
| English | n = 8 | IntHigh | 91.7% | 70.8% | 50% | 95.8% |
| English | n = 5 | Advanced | 100% | 93.3% | 66.6% | 100% |
| Control | n = 26 | L1 | 93.6% | 94.9% | 94.9% | 97.4% |

When short-distance and long-distance were compared, the correct response rate for the LD sentences was significantly lower than that for the short-distance sentences within each of the Chinese and the English groups. Chinese-speaking JSL learners and English-speaking JFL learners did not show as great an understanding of long-distance *zibun* binding as the native speakers did. Moreover, the Chinese-speakers showed some significant improvement in their understanding of long-distance *zibun* binding as they stayed longer in Japan. On the other hand, the English speakers failed to show such improvement in their understanding of long-distance *zibun* binding. These results constitute empirical evidence that the long-distance *zibun* binding is indeed difficult crosslinguistically for both Chinese and English speakers of L2 Japanese. They also suggest that L1 knowledge may to some degree function to help L2 learners understand long-distance reflexive binding. Each L2 group behaves in a

similar way as the L1 group in short-distance binding, but in a different way in long-distance binding from the L1 group.[33]

Yoshimura et al. (2013) show the acquisition of L2 and L3 Japanese *zibun* by L1 Chinese, English, and Turkish speakers, employing the truth value judgment task with similar test stimuli as above (Chinese learners of L3 Japanese (L3C), Chinese learners of L2 Japanese (L2C), English-speaking learners of L2 Japanese (L2E), Turkish learners of L3 Japanese (L3T)) and Control groups participated in their study with a truth judgment task. Test situations and sentences are similar to those in (40) above. Table 10 shows a summary of the results.

**Table 10:** Correct response rates by group and sentence type

| | | | Short-distance | | Long-distance | |
|---|---|---|---|---|---|---|
| L1 | Participant # | Japanese | TRUE | FALSE | TRUE | FALSE |
| Chinese | n = 18 | L2 IntMid | 72.2% | 87.0% | 66.7% | 77.8% |
| Chinese | n = 30 | L3 IntMid | 71.7% | 97.0% | 80.0% | 97.8% |
| Turkish | n = 40 | L3 IntMid | 84.2% | 70.0% | 59.2% | 82.5% |
| English | n = 13 | L2 IntHigh | 92.3% | 79.5% | 56.4% | 91.4% |
| Control | n = 26 | L1 | 94.2% | 94.9% | 93.6% | 96.1% |

The results on True sentences evoked significant main effects both on participant groups and sentence type (short-distance vs. long-distance) as well as a significant interaction between groups and types. A post-hoc test indicated that the L3C group was significantly lower than the Control group in the short-distance condition, and L2C, L2E, and L3T groups all evoked fewer correct responses than the Control group in the long-distance condition. This suggests that L1 transfer did not seem to have occurred in the acquisition of *zibun*. The results show that L2 and L3 learners all have acquired adequate sensitivity to the locality requirement in the interpretation of *zibun*. Locality is assumed to be the core notion underlying language acquisition, and thus short-distance binding did not pose any serious problem for L2 and L3 learners with different L1s. Because *zibun*, *ziji* and *kendi* may become logophoric pronouns in the discourse context (e.g., Kuno 1976; Huang and Tang 1991; Demirci 2001), one would think that Chinese and Turkish speakers would transfer their knowledge of logophoric pronouns to L2 or L3 Japanese. But that was not the case, although L3C acquired long-distance *zibun* earlier than English-speaking learners.[34]

**33** Note that the English speakers' mean correct response rate of 56.4% in the present study is much higher than the high proficiency English group's long-distance acceptance rate of 7.7% in Thomas (1991). We speculate that this discrepancy comes from the task difference.

**34** Based on Yuan's (1998) analysis of Japanese speakers' ease with the long-distance binding of the Chinese *ziji*, we suggest that similarities between *ziji* and *zibun* can function to facilitate the Chinese speakers' development of long-distance binding in Japanese, especially L3. Our view that L1 pragmatics may be a facilitator in L2 learning at the syntax-pragmatics interface needs to be further explored. See Thomas (1991, 1993) for a discussion of how pragmatic effects may or may not affect both native- and non-native speakers' preference for subject-orientation in the interpretation of *himself*. See also Sachs (2010) on individual differences.

The short-distance – long-distance asymmetry can be accounted for and these results constitute empirical evidence in support of the modularity of syntax and pragmatics in L2/L3 acquisition.[35]

The parameter resetting approach adopted in almost all previous studies on the acquisition of anaphor binding cannot furnish a plausible account for the cross-linguistic short-distance – long-distance asymmetry because a long-distance anaphor, the counterpart of *zibun*, does exist in Chinese and Turkish, but not in English: The parameter resetting hypothesis predicts that L1 Chinese and Turkish speakers should have gained a grammatical knowledge of long-distance binding earlier than L1 English speakers. However, the prediction is not borne out. We argue that, being anaphoric, the short-distance *zibun* is subject to Binding Principle A, so the locality condition follows, while the long-distance *zibun* requires L2/3 learners to understand a relevant pragmatic principle of aboutness beyond syntax. The complexity of the pragmatic knowledge involved in the notion of logophoricity needs time for L2/3 learners to capture.

# 4 Concluding remarks

We have discussed the modularity of syntax, morpho-syntax, syntax-semantics/pragmatics interfaces referring to Japanese speakers' L2 acquisition of English subjects/expletives, WH-interrogatives, inflectional morphology, bound variable interpretations, and reflexives, English speakers' L2 acquisition of Japanese bound variable interpretations, and *zibun*, and Chinese and Turkish speakers' L2 acquisition of *zibun*. By discussing them, the idea that core or "narrow" syntactic properties are acquired early while the acquisition of interface (syntax-morphology/phonology, syntax-semantics/pragmatics) properties is delayed in L2 acquisition was supported. According to this approach, the fact that different grammatical properties are acquired at different rates despite the L1 and L2 differences can be explained rather straightforwardly. Although this theoretical approach is more promising, more research on Japanese grammatical properties that fall in interfaces is necessary. While investigating them, it is also important to consider L2 learners' performance limitations (i.e., computational limitation). See Sawasaki and Kashiwagi's chapter in this volume.

# Acknowledgments

We would like to thank the participants of the Fall 2012 Japanese Language Processing class, Akiko Kashiwagi-Wood, and Seth Wiener for their helpful comments on

**35** The modularity of syntax and pragmatics is argued for in Chien and Wexler (1990) who found the early acquisition of Binding A but the delayed acquisition of Binding B among L1 English-speaking children.

the earlier versions of this chapter. We also acknowledge that the first author has been supported by the National Institute for Japanese Language and Linguistics, Tokyo, Japan, and the Department of East Asian Languages and Literatures and the Institute for Japanese Studies, The Ohio State University and the second author has been supported by the University of Shizuoka. Many of the research projects cited in this chapter have been supported by the Grant-in-Aid for Scientific Research (C) 20520551 and (B) 233020116 from the Japan Society for the Promotion of Science to the second author. Their generous support is gratefully acknowledged.

## References

Abe, Jun. 1997. The locality of *zibun* and logophoricity. *Researching and Verifying an Advanced Theory of Human Language, COE Research Report*, 595–626. Kanda University of International Studies.

Aoshima, Sachiko, Colin Phillips and Amy Weinberg. 2004. Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language* 51. 23–54.

Bard, Ellen Gurman, Dan Robertson and Antonella Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72(1). 32–68.

Beck, Maria-Luise. 1998. L2 acquisition and obligatory head movement: English-speaking learners of German and the Local Impairment Hypothesis. *Studies in Second Language Acquisition* 20. 311–348.

Bley-Vroman, Robert. 1989. What is the logical problem of foreign language learning? In Susan M. Gass and Jacquelyn Schachter (eds.), *Linguistic perspectives on second language acquisition*, 41–68. Cambridge: Cambridge University Press.

Chien, Yu-Chin and Kenneth Wexler. 1990. Children's knowledge of locality condition in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1. 225–295.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson and Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1981. *Lectures on government and binding.* Dortrecht: Foris.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin (ed.), *Principles and parameters in comparative grammar*, 417–454. Cambridge, MA: MIT Press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Ken Hale and Samuel J. Keyser (eds.), *The view from Building 20*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2007a. Approaching UG from below. In Uli Sauerland and Hans-Martin Gärtner (eds.), *Interfaces+Recursion=Language?*, 1–29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2007b. On phases. In Robert Freidin, Carlos Otero and Maria-Luisa Zubizarreta (eds.), *Foundational issues in Linguistic theory*, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2007c. Biolinguistic explorations: Design, development, evolution. *International Journal of Philosophical Studies* 15. 121. London: Routledge.

Clahsen, Harald and Pieter Muysken. 1989. The UG paradox in L2 acquisition. *Second Language Research* 5. 1–29.

Clements, George N. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages* 2. 355–406.

Crain, Stephen and Cecile McKee. 1986. Acquisition of structural restrictions on anaphora. *Proceedings of the North Eastern Linguistic Society 16*, 94–110. Amherst, MA: University of Massachusetts.
Demirci, Mahide. 2001. Acquisition of binding of English reflexives by Turkish L2 learners: A Neo-Gricean pragmatic account. *Journal of Pragmatics* 33. 753–775.
Eckman, Fred R. 1996. On evaluating arguments for special nativism in second language acquisition theory. *Second Language Research* 12(4). 398–419.
Embick, David and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32. 555–595.
Embick, David and Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39. 1–53.
Finer, Daniel L. and Ellen L. Broselow. 1986. Second language acquisition of reflexive binding. *Proceedings of the North Eastern Linguistic Society 16*, 154–168. Amherst, MA: University of Massachusetts.
Goad, Heather and Lydia White. 2004. Ultimate attainment of L2 inflection: Effects of L1 prosodic structure. In Susan Foster-Cohen (ed.), *EUROSLA Yearbook* 4, 119–145. Amsterdam: John Benjamins.
Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale and Samuel J. Keyser (eds), *The view from Building 20: Essays in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.
Hamilton, Robert. 1996. Against underdetermined reflexive binding. *Second Language Research* 12(4). 420–446.
Hawkins, Roger. 2005. Explaining full and partial success in the acquisition of second language grammatical properties. *Second Language* 4. 7–25.
Hawkins, Roger and Cecilia Y.-H. Chan. 1997. The partial availability of Universal Grammar in second language acquisition: The Failed Functional Features Hypothesis. *Second Language Research* 13. 187–226.
Hawkins, Roger and Hajime Hattori. 2006. Interpretation of English multiple Wh-questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research* 22(3). 269–301.
Hawkins, Roger and Sarah Liszka. 2003. Locating the source of defective past tense marking in advanced L2 English speakers. In Roeland van Hout, Aafke Hulk, Folkert Kuiken and Richard Towell (eds.), *The lexicon-syntax interface in second language acquisition*, 21–44. Amsterdam: John Benjamins.
Haznedar, Belma. 2001. The acquisition of the IP system in child L2 English. *Studies in Second Language Acquisition* 23. 1–39.
Haznedar, Belma and Bonnie D. Schwartz. 1997. Are there optimal infinitives in child L2 acquisition? In Elizabeth Hughes, Mary Hughes, and Annabel Greenhill (eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development*, 257–268. Somerville, MA: Cascadilla Press.
Hirakawa, Makiko. 1989. The government category parameter in second language acquisition. In Yukio Otsu (ed.), *MITA working papers in psycholinguistics 2*, 27–40.
Hirakawa, Makiko. 1990. A study of the L2 acquisition of English anaphors. *Second Language Research* 6. 60–85.
Hoji, Hajime. 1991. Kare. In Carol Georgopoulos and Roberta Ishihara (eds.), *Interdisciplinary approaches to language*, 287–304. Dordrecht: Kluwer Academic Publishers.
Huang, C.-T. James and C.-S. Luther Liu. 2001. Logophoricity, attitudes and *ziji* at the interface. In Peter Cole, Gabriella Hermon and C.-T. James Huang (eds.), *Long distance reflexives*, 141–195. New York: Academic Press. C.-S. L.
Huang, C.-T. James and C.-C. Jane Tang. 1991. The local nature of the long-distance reflexive in Chinese. In Jan Koster and Eric Reuland (eds.), *Long-distance anaphora*, 263–282. Cambridge: Cambridge University Press.

Ionin, Tania and Kenneth Wexler. 2002. Why is 'is' easier than '-s'?: Acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research* 18. 95–136.

Ito, Miyoko. 2003. The interpretation of pronouns by Japanese learners of English. In Susan H. Foster-Cohen and Simona Pekarek Doehler (eds.), *EUROSLA Yearbook 3*, 29–56. Amsterdam: John Benjamins.

Kanagy, Ruth. 1994. Developmental sequences in learning Japanese: A look at negation. *Issues in Applied Linguistics* 5(2). 255–277.

Kaneko, Asuka. 2005. *Successive Cyclic Wh-movement in Second Language Acquisition*. Shizuoka: University of Shizuoka MA thesis.

Kanno, Kazue. 1997. The acquisition of null and overt pronominals in Japanese by English speakers. *Second Language Acquisition* 5. 317–332.

Kanno, Kazue. 1998. The stability of UG principles in second language acquisition: Evidence from Japanese. *Linguistics* 36(6). 1125–1146.

Kano, Akihiro and Mineharu Nakayama. 2004a. Knowledge of binding and the role of empathy in interpreting anaphora among adult second-language learners. In Masahiko Minami, Harumi Kobayashi, Mineharu Nakayama and Hidetosi Sirai (eds.), *Studies in language sciences 3*, 169–184. Tokyo: Kurosio Publishers.

Kano, Akihiro, and Mineharu Nakayama. 2004b. Variable binding and *zibun* in L2 Japanese. *Ars Linguistica* 11. 41–67.

Kim, Sun-Hee. 1992. Division of labor between grammar and pragmatics concerning anaphora. *Kansas Working Papers in Linguistics* 17. 191–221.

Kuribara, Chieko. 2003. Subjects and verbal inflections in SLA: In defense of "full transfer/limited access" model. *Gaikokugo Kyōiku Kenkyu* 5. 17–40. (Osaka: Kansai University)

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kuno, Susumu. 1976. Subject, theme, and the speaker's empathy: A re-examination of reflexivization phenomena. In Charles Li (ed.), *Subject and topic*, 417–444. New York: Academic Press.

Kuroda, Shige-Yuki. 1973. On Kuno's direct discourse analysis of the Japanese reflexive *zibun*. *Papers in Japanese Linguistics* 2. 136–147.

Kuroda, Shige-Yuki. 1988. Whether we agree or not: A comparative syntax of English and Japanese. *Linguisticae Investigationes* 12. 1–47.

Lardiere, Donna. 1998a. Case and tense in the 'fossilized' steady state. *Second Language Research* 14. 1–26.

Lardiere, Donna. 1998b. Dissociating syntax from morphology in a divergent end-state grammar. *Second Language Research* 14. 359–375.

Lardiere, Donna. 2000. Mapping features to forms in second language acquisition. In John Archibald (ed.), *Second language acquisition and linguistic theory*, 102–129. London: Blackwell.

Lardiere, Donna. 2005. On morphological competence. In Laurent Dekydtspotter, Rex A. Sprouse and Audrey Liljestrand (eds.), *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)*, 178–192. Somerville, MA: Cascadilla Proceedings Project.

Lasnik, Howard. 1995. Verbal morphology: Syntactic Structures meets the Minimalist Program. In Campos, Héctor and Paula Kempchinsky (eds.), *Evolution and revolution in linguistic theory*, 251–275. Washington, DC: Georgetown University Press. (Also in Lasnik, Howard. 1999. *Minimalist analysis*, 97–119. Oxford: Blackwell.)

Lieberman, Moti, Sachiko Aoshima, and Colin Phillips. 2006. Native-like biases in generation of Wh-questions by non-native speakers of Japanese. *Studies in Second Language Acquisition* 28. 423–448.

Masumoto, Ayaka. 2008. *Overt pronouns and bound variable reading in L2 Japanese*. Columbus, OH: The Ohio State University MA thesis.

Manzini, M. Rita and Kenneth Wexler. 1987. Parameters, binding theory, and learnability. *Linguistic Inquiry* 18. 413–444.

Matsumura, Masanori. 2007. Semantic behind the structure, and how it affects the learner: A new perspective on second language reflexives. *IRAL*. 321–352.

Miyamoto, Yoichi, and Keiko Okada. 2004. Topicalization and WH-movement in the grammar of Japanese EFL learners. In Alejna Brugos, Liunea Micciulla and Christine E. Smith (eds), *The proceedings of the 24th annual Boston University conference on language development*, 375–389. Somerville, MA: Cascadilla Press.

McCloskey, James. 1991. There, it, and agreement. *Linguistic Inquiry* 22. 563–567.

Montalbetti, Mario M. 1984. *After binding: On the interpretation of pronouns*. Cambridge, MA: Massachusetts Institute for Technology dissertation.

Murasugi, Keiko. 1991. *Noun phrase in Japanese and English: A study in syntax, learnability and acquisition*. Storrs, CT: University of Connecticut dissertation.

Murasugi, Keiko. 2015. Root infinitive analogues in Child Japanese. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: Mouton de Gruyter.

Nakahama, Yuko. 2011. *Referent markings in L2 narratives*. Tokyo: Hituzi Syobo.

Nakayama, Mineharu and Akihiro Kano. 2007. JFL learners' interpretations of *zibun* and *zibuntachi*. In Masahiko Minami (eds.), *Applying theory and research to learning Japanese as a foreign language*, 113–129. Cambridge: Cambridge Scholars Press.

Nakayama, Mineharu, Shigenori Wakabayashi and Hironobu Hosoi. 2006. Japanese EFL learners' interpretations of possessive pronouns. Paper presented at EUROSLA 16, Boğaziçi Üniversitesi, Antalya, 14 September.

Noguchi, Tohru. 1997. Type types of pronouns and variable binding. *Language* 73(4). 770–797.

O' Grady, William. 1996. Language acquisition without Universal Grammar: A general nativist proposal for L2 learning. *Second Language Research* 12(4). 374–397.

Osterhout, Lee and Kayo Inoue. 2007. What the brain's electrical activity can tell us about language processing and language learning. In Tsutomu Sakamoto (ed.) *Communicating skills of intention*, 293–309. Tokyo: Hituzi Shobo.

Oya, Shinobu. 2006. *The acquisition of bound variables by Japanese EFL learners*. Shizuoka: University of Shizuoka MA thesis.

Phinney, Marianne. 1987. The pro-drop Parameter in Second Language Acquisition. In Thomas Roeper and Edwin Williams (eds.), *Parameter setting*, 221–238. Dordrecht: Reidel Publishing Company.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In Eric J. Reuland and Alice G. B. ter Meulen (eds.), *The representation of (in)definiteness*. 98–129. Cambridge, MA: MIT Press.

Pérez-Leroux, Ana T. and William R. Glass. 1999. Null anaphora in Spanish second language acquisition: Probabilistic versus generative approaches. *Second Language Research* 15(2). 220–249.

Pimentel, Carlos L. and Mineharu Nakayama. 2012a. Pronominal interpretations in L2 Japanese. *Journal of Japanese Linguistics* 28, 111–131.

Pimentel, Carlos L. and Mineharu Nakayama. 2012b. L2 pronominal interpretations revisited. Poster presented at the 14th annual meeting of the Japan Society of Language Sciences, Nagoya University, 30 June.

Prévost, Philippe. 2008a. Morphological variability in the development of L2 French morphosyntax: The issue of impairment and L1 influence. In Roumyana Slabakova, Silvina A. Montrul and Philippe Prévost (eds.), *Issues in linguistic development: In honor of Lydia White*, 135–156. Amsterdam: John Benjamins.

Prévost, Philippe. 2008b. Knowledge of morphology and syntax in early adult L2 French: Evidence for the Missing Surface Inflection Hypothesis. In Juana M. Liceras, Helmut Zobl and Helen Goodluck (eds.), *The role of formal features in second language acquisition*, 352–377. New York: Lawrence Erlbaum Associates.

Prévost, Philippe and Lydia White. 2000a. Accounting for morphological variation in second language acquisition: Truncation or missing inflection? In Luigi Rizzi and Marc-Ariel Friedemann (eds.), *The acquisition of syntax: Issues on comparative developmental linguistics*, 202–235. London: Longman.

Prévost, Philippe and Lydia White. 2000b. Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16. 103–133.

Radford, Andrew. 1997. *Syntactic theory and the structure of English*. Cambridge: Cambridge University Press.

Radford, Andrew. 2004. *English syntax: An Introduction*. Cambridge: Cambridge University Press.

Radford, Andrew. 2009. *An introduction to English sentence structure*. Cambridge: Cambridge University Press.

Radford, Andrew and Hideki Yokota. 2006. UG-constrained Wh-movement in Japanese learners' English questions. *Second Language* 5. 61–94

Reinhart, Tanya. 2006. *Interface strategies: optional and costly computations*. Cambridge, MA: MIT Press.

Sachs, Rebecca R. 2010. *Individual differences and the effectiveness of visual feedback on reflexive binding in L2 Japanese*. Georgetown, DC: Georgetown University dissertation.

Safir, Ken. 1985. *Syntactic chains*. Cambridge: Cambridge University Press.

Saito, Mamoru and Hajime Hoji. 1983. Weak crossover and move alpha in Japanese. *Natural Language and Linguistic Theory* 1&2. 245–259.

Sano, Tetsuya. 2002. *Roots in language acquisition: A comparative study of Japanese and European languages*. Tokyo: Hituzi Syobo.

Sawasaki, Koichi and Akiko Kashiwagi-Wood. 2015. Issues in L2 Japanese sentence processing: similarities/differences with L1 and individual differences in working memory. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: Mouton de Gruyter.

Schwartz, Bonnie and Rex Sprouse. 1994. Word order and nominative case in non-native language acquisition: A longitudinal study of German interlanguage. In Teun Hoekstra and Bonnie Schwartz (eds.), *Language acquisition studies in generative grammar*, 317–368. Amsterdam: John Benjamins.

Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18. 445–479.

Sheen, Ron. 2000. A response to Kanno's "The stability of UG principles in second-language acquisition: Evidence from Japanese". *Linguistics* 38(4). 799–816.

Shibuya, Mayumi and Shigenori Wakabayashi. 2008. Why are L2 learners not always sensitive to subject-verb agreement? In Leah Roberts, Florence Myles and Annabelle David (eds.) *EUROSLA Yearbook* 8, 235–258. Amsterdam: John Benjamins.

Shibuya, Mayumi, Shigenori Wakabayashi and Tae Yamazaki-Hasegawa. 2009. Overuse of English 3rd person singular *-s* in production and perception among Japanese learners. Poster presented at the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009), University of Illinois, Urbana-Champaign, 14 March.

Shirahata, Tomohiko. 2002. The acquisition of Japanese binding form, *zibun* by English-speaking children. *Second Language* 1. 62–96.

Shirahata, Tomohiko. 2007. Interpretation of English pronouns and reflexives by Japanese learners. *Shizuoka University School of Education research report* (Humanities, Social Sciences) 57. 141–156.

Shirahata, Tomohiko and Junko Ishigaki. 2001. The acquisition of subject-orientation and long distance binding in *zibun* by Japanese L1 children. *Ars Linguistica* 1. 41–59.

Shirai, Yasuhiro. 2015. The L2 acquisition of Japanese. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: Mouton de Gruyter.

Sorace, Antonella. 2007. Optionality at the syntax-discourse interface in near-native L2 speakers. *Second Language* 6. 3–15.
Sorace, Antonella and Francesca Filiaci. 2006. Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22. 339–368.
Suda, Koji and Shigenori Wakabayashi. 2007. The acquisition of pronominal case- marking by Japanese learners of English. *Second Language Research* 23. 179–214.
Slabakova, Roumyana. 2009. What is easy and what is hard to acquire in a second language? In Melissa Bowles, Tania Ionin, Silvina Montrul, and Annie Tremblay (eds.), *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*, 280–294. Somerville, MA: Cascadilla Proceedings Project.
Sturt, Patrick. 2003. A new look at the syntax-discourse interface: The use of binding principles in sentence processing *Journal of Psycholinguistic Research* 32(2). 125–139.
Thomas, Margaret. 1991. Universal grammar and the interpretation of reflexives in a second language. *Language* 67. 211–239.
Thomas, Margaret. 1993. *Knowledge of reflexives in a second language*. Philadelphia: John Benjamins.
Thomas, Margaret. 1995. Acquisition of the Japanese reflexive *zibun* and movement of anaphors in logical form. *Second Language Research* 11. 206–234.
Thomas, Margaret. 2006. Japanese, the grammar of reflexives, and second language acquisition. In Mineharu Nakayama, Reiko Mazuka and Yasuhiro Shirai (eds.), *Handbook of East Asian psycholinguistics*, Vol. 2, *Japanese*. Cambridge: Cambridge University Press.
Umeda, Mari. 2005. Wh-movement in Japanese-English Interlanguage. In Alejna Brugos, Manuella R. Clark-Cotton, and Seungwan Ha (eds.), *Proceedings of the 29th annual Boston University Conference on Language Development*, 616–626. Sommerville, MA: Cascadilla.
Wakabayashi, Shigenori. 1996. The nature of SLA: SLA of English reflexives. *Second Language Research* 12. 266–303.
Wakabayashi, Shigenori. 1997. The acquisition of functional categories by learners of English. Cambridge: University of Cambridge dissertation.
Wakabayashi, Shigenori. 2002. The acquisition of null subjects in English: A minimalist account. *Second Language Research* 18. 28–71.
Wakabayashi, Shigenori and Izumi Okawara. 2003. Japanese learners' errors in long distance Wh-questions. In Shigenori Wakabayashi (ed.) *Generative approaches to the acquisition of English by native speakers of Japanese*, 215–245. Berlin: Mouton de Gruyter.
Wakabayashi, Shigenori, Kazuhiko Fukuda, Masanori Bannai and Shoichi Asaoka. 2007. Japanese speakers' sensitivity to third person singular *–s* in English: Arguments based on ERP data. *Second Language* 6. 19–46.
Wakabayashi, Shigenori and Tae Yamazaki. 2006. Santangen-no *–s* no ayamari-ni mirareru tōgokōzō-to senteki kyori-no eikyō [Effects of the syntactic distance and of the linear distance in errors of 3rd person singular *–s*]. *Kagaku-kenkyū-hi hōkokusho* 15520364 [Kakenhi Technical Report 15520364], 45–64.
Watanabe, Eriko, Chisato Fuji, Yoshie Kabuto and Keiko Murasugi. 2008. Experimental evidence for the parameter resetting hypothesis: The second Language acquisition of English reflexive-binding by Japanese speakers. *Nanzan Linguistics: Special Issue* 3(2). 263–283.
Wexler, Kenneth and Rita M. Manzini. 1987. Parameters and learnability in binding theory. In Thomas Roeper and Edwin Williams (eds.), *Parameter setting*, 41–76. Dordrecht: D. Reidel Publishing Company.
White, Lydia. 1989. *Universal Grammar and second language acquisition*. Philadelphia: John Benjamins.
White, Lydia. 2000. Second language acquisition: From initial to final state. In John Archibald (ed.), *Second language acquisition and linguistic theory*, 130–155. Malden, MA: Blackwell.

White, Lydia. 2003a. Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition* 6. 129–141.
White, Lydia. 2003b. *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
White, Lydia. 2008. Some puzzling features of L2 features. In Juana M. Liceras, Helmut Zobl and Helen Goodluck (eds.), *The role of formal features in second language acquisition*, 299–326. New York: Lawrence Erlbaum Associates.
White, Lydia. 2011. Second language acquisition at the interfaces. *Lingua* 121. 577–590.
Wolfe-Quintero, Kate. 1996. Nativism does not equal Universal Grammar. *Second Language Research* 12(4). 335–373.
Yamashita, Yuka. 2007. *The Acquisition of pied-piping and preposition stranding by Japanese EFL learners*. Shizuoka: University of Shizuoka M.A. thesis.
Yoshimura, Noriko and Mineharu Nakayama. 2009. Nominative case marking and verb inflection in L2 grammar: Evidence from Japanese college students' compositions. *The Proceedings of the Tenth Tokyo Conference on Psycholinguistics*, 359–383. Tokyo: Hituzi Syobo.
Yoshimura, Noriko and Mineharu Nakayama. 2010a. *Kaigai tanki eigo kenshū-to dai 2 gengo shūtoku* [Short-term English study abroad and L2 acquisition]. Tokyo: Hituzi Syobo.
Yoshimura, Noriko and Mineharu Nakayama. 2010b. Expletives in L2 English and narrow syntax. *Ars Linguistica* 17. 161–175.
Yoshimura, Noriko and Mineharu Nakayama. 2011. L2 acquisition of overt WH-movement revisited. *Ars Linguistica* 18. 194–216.
Yoshimura, Noriko, Mineharu Nakayama, Tomohiko Shirahata, Koichi Sawasaki and Yasuo Terao. 2012. *Zibun* and locality in L2 Japanese. *Journal of Japanese Linguistics* 28. 91–112.
Yoshimura, Noriko, Mineharu Nakayama, Koichi Sawasaki, Atsushi Fujimori and Hiroya Shimizu. 2012. L2 knowledge at the syntax-pragmatics interface: Interpretations of reflexives by Japanese, Korean, and Chinese ESL learners. In Yukio Otsu (ed.), *The Proceedings of the Thirteenth Tokyo Conference on Psycholinguistics*, 303–323. Tokyo: Hituzi Syobo.
Yoshimura, Noriko, Mineharu Nakayama, Koichi Sawasaki, Atsushi Fujimori and Bariş Kahraman. 2013. The development of long-distance *zibun*: Roles of L1 and L2 in L3 acquisition. In Yukio Otsu (ed.), *The Proceedings of the Fourteenth Tokyo Conference on Psycholinguistics*, 221–236. Tokyo: Hituzi Syobo.
Yoshimura, Noriko, Mineharu Nakayama, Koichi Sawasaki, and Hiroya Shimizu. 2013. Locality in L2 Japanese and English. In Stavroula Stavrakaki, Marina Lalioti, and Polyxeni Konstantinopoulou (eds.), *Advances in language acquisition*, 375–383. Newcastle upon Tyne: Cambridge Scholars Publishing.
Yuan, Boping. 1998. Interpretation of binding and orientation of the Chinese reflexive *ziji* by English and Japanese speakers. *Second Language Research* 14. 324–340.

Alison Gabriele and Mamori Sugita Hughes

# 9 Tense and aspect in Japanese as a second language

## 1 Introduction

Tense and aspect have long been of interest to researchers examining second language (L2) acquisition in a range of languages, including Japanese. This interest is due in large part to the complexity of these linguistic phenomena and the difficulty of understanding the boundary between the two domains. Tense and aspect present a challenge because the learner needs to discern the precise contribution of each lexical and grammatical element to the overall interpretation and crucially understand the interaction between them. It is the complex nature of the mapping between form and meaning in the domain of tense and aspect that has inspired many studies on language acquisition. In the past decade, research on the acquisition of Japanese as a second language in particular has pushed this area of research even further as studies have examined a wider array of factors, including the role of input and transfer from the native language (L1), that may impact development and ultimate attainment in this domain (see Shirai, this volume, for a general overview on second language acquisition research in Japanese).

We will begin by defining the relevant terms. Comrie (1976) distinguished tense and aspect by proposing that "Tense relates the time of the situation referred to to some other time, usually to the moment of speaking" (Comrie 1976: 1–2) while "aspects are different ways of viewing the internal temporal constituency of a situation" (Comrie 1976: 3). In other words, tense locates an event on a timeline in relation to the moment of speech while aspect looks into the internal structure of a particular situation or event, focusing on how an event unfolds. Many languages, including Japanese, encode the aspectual distinction between *perfective* and *imperfective*. In Comrie's terms, perfective aspect views a situation as a whole from the outside without distinguishing any internal structure, while imperfective aspect, in contrast, views the internal structure of a situation, disregarding the beginning and endpoint of an event (Comrie 1976; Smith 1991/1997). Aspect is also encoded in the properties of the verb phrase: *lexical aspect* generally refers to Vendler's four-way classification, which distinguishes states such as *know*, activities such as *walk*, accomplishments such as *draw a circle*, and achievements such as *die* (Vendler 1967). Although there is specific discussion of the lexical aspectual classes in Japanese (Jacobsen 1992; Kindaichi 1950; McClure 1995), we will rely on Vendler's description as it is the one used in most studies of second language acquisition. Vendler (1967) distinguished the aspectual classes on the basis of semantic features, as is shown in

Table 1. The semantic features underlying the aspectual classes are available in every language and are argued to be universals (Olsen 1997; Van Valin 2006; Von Fintel and Matthewson 2008).

**Table 1:** Vendler's (1967) aspectual classes as defined by semantic features

| | Telic | Dynamic | Durative |
|---|---|---|---|
| State | – | – | + |
| Activity | – | + | + |
| Accomplishment | + | + | + |
| Achievement | + | + | – |

The feature +/-telic encodes whether or not a verb phrase specifies an inherent endpoint. Accomplishments and achievements both entail a change of state and thus encode an endpoint, while states and activities do not. Accomplishments and achievements are themselves distinguished by the feature +/-durative, which refers to whether or not the verb phrase encodes a process. Achievements, which are said to occur instantaneously, do not encode a process. The feature +/-dynamic distinguishes states from the other classes.

Klein (2009) described what he considered to be an idealized system for tense and aspect in which a particular language would have one set of dedicated tense morphemes to encode past, present, and future and another set of dedicated aspect markers that would encode for example whether an event is in progress or complete. In this idealized system, the morphemes would combine freely so that the encoding of notions such as *ongoing in the past* would be morphologically transparent, combining a past tense morpheme and an ongoing aspect morpheme. Although there are examples of such transparency in natural languages, in large part the mapping between form and meaning is far more intricate than an idealized system might allow. Languages differ with respect to whether tense and aspect are overtly realized in the grammatical system (for example, encoded in the morphology) or whether temporal or aspectual notions are expressed via other means such as temporal adverbs (*yesterday*). In languages such as Mandarin Chinese or Thai, tense is not encoded grammatically but rather is expressed via other means such as aspectual markers or adverbs (Li and Thompson 1981; Smith 2008). In languages such as German and Inuktitut, aspect is not overtly realized in the grammatical system. In these languages, aspect is expressed only via lexical aspect, encoded in the semantic properties of the verb phrase (Bohnemeyer and Swift 2004). Thus, within individual languages, it is clear that temporal and aspectual notions are truly compositional in that they rely on a complex interaction of grammatical markers, the properties of the verb phrase, and other elements at the sentence level such as adverbial phrases (Olsen 1997; Smith 1991/1997; Verkuyl 1993). The precise nature of this interaction differs across languages depending on the grammatical categories

that are represented and the specific properties of those grammatical forms. These crosslinguistic differences present a very interesting test case in second language acquisition as researchers are interested in the extent to which properties of the learners' native language impact development in the L2.

The question of how temporal and aspectual notions are encoded in the grammar of the learner is as central to the study of language acquisition as it is to the study of the linguistic system itself. A large body of research in second language acquisition examines what has been called the Aspect Hypothesis (Andersen and Shirai 1996; Bardovi-Harlig and Bergström 1996; Shirai 1991), which predicts that learners are limited with respect to their distribution of temporal and aspectual forms. Learners tend to use past or perfective forms with telic verb phrases such as accomplishments and achievements and use present or imperfective forms with activity verbs. The proposal is that learners at early stages align temporal and aspectual markers with specific lexical aspectual features. Telic verb phrases, which encode an endpoint, are aligned with perfective aspectual forms, which help to define that endpoint by encoding completion, while atelic verb phrases such as activities are aligned with imperfective forms which help to define the durativity. Researchers have examined whether this pattern, in which 'prototypical' associations (Shirai and Andersen 1994) emerge first in development, holds across learners of several different languages. The question is whether L2 learners generally follow similar developmental paths, as is proposed by the Aspect Hypothesis, or whether development is influenced by the properties of the learner's native language.

Several recent L2 studies on the acquisition of tense and aspect in Japanese have focused on the issue of transfer both to better understand the patterns observed in research on the Aspect Hypothesis (Sugaya and Shirai 2007) and to shed light on several other theoretical issues in the domain of L2 acquisition. To this end, transfer has been examined at various "levels" or domains within the broader categories of tense and aspect including grammatical aspect, lexical aspect, and tense. This interest in transfer, a central issue in the study of L2 acquisition, spans across different theoretical perspectives, including language acquisition researchers who approach the study of tense and aspect from both functional (e.g. Sugaya and Shirai 2007) and generative (e.g Gabriele 2009) frameworks.

This recent research has examined the extent to which specific learnability scenarios either facilitate or impede learners' ability to overcome transfer (Gabriele 2005, 2009, 2010) and whether properties which are not instantiated in the learners' L1 can be successfully acquired in the L2 (Gabriele and McClure 2011). Recent work has also examined whether transfer is more prevalent at the level of lexical aspect than at the level of grammatical aspect in order to better understand what linguistic properties are candidates for transfer in L2 acquisition (Nishi 2008, see also Gabriele, McClure, and Martohardjono 2003). Finally, one very recent study examines the extent to which the properties of the learners' L1 impact L2 processing in the domain of tense and aspect (Long, Nakaishi, Ono, and Sakai 2012). This work examines whether L2

learners can use temporal and aspectual information in the course of online sentence processing similarly to native speakers, extending work in this domain into the realm of psycholinguistics.

This chapter reviews this recent work to better understand in which domains there is evidence for transfer, the extent to which L1 influence can be overcome at higher levels of proficiency, and how transfer interacts with other factors, such as the universal predispositions argued for in the Aspect Hypothesis, and the specific linguistic properties of the target language. The overall picture that emerges is that a complex interplay of factors is at work: both transfer at the level of grammatical aspect and the specific linguistic encoding of tense and aspect in the target language are important determinants of L2 development and ultimate attainment in Japanese.

In the following section we review the relevant linguistic facts on tense and aspect in Japanese. In Section 3 we review several recent studies that have examined the Aspect Hypothesis in L2 Japanese, focusing on papers published subsequent to Shirai's (2002a) review of this literature. In Section 4 we review some recent studies that have examined transfer in L2 Japanese at the levels of grammatical aspect, lexical aspect, and within the noun phrase. In Section 5 we review a very recent study of the processing of tense and aspect in L2 Japanese, and in Section 6 we draw general conclusions about recent research in this domain and point out some future directions for research.

## 2 Tense and aspect in Japanese

All sentences in Japanese are obligatorily marked for tense (past or non-past). As is shown in (1), with dynamic verb phrases, such as activities, the morpheme –(*r*)*u* encodes either a habitual reading or future tense. In the absence of adverbs (*often*, *tomorrow*), the specific reading will be determined by context. Japanese does not have an independent morphological marker of future tense, such as the auxiliary *will* in English. With stative verbs, as in (2), a verb inflected with the non-past morpheme refers to a present state.

(1) *Tomoko wa* *susi o* *tabe-<u>ru</u>.*
Tomoko TOP sushi ACC eat-NONPST
'Tomoko eats/will eat sushi.'

(2) *Tomoko wa* *tookyoo ni* *i-<u>ru</u>.*
Tomoko TOP Tokyo LOC be-NONPST
'Tomoko is in Tokyo.'

The morpheme *–ta* encodes past tense in (3) and can be attached to verbs from any of the four lexical aspectual classes. There is debate in the literature with respect to whether *–ta* is a past tense marker or a marker of perfect aspect (see review in Ogihara 1998), which Shirai (2002: 43) proposes is due to the fact that the marker is in the process of grammaticizing from a perfect marker to a marker of simple past tense.

(3) *Tomoko wa* *susi o* *tabe-ta.*
Tomoko TOP sushi ACC eat-PST
'Tomoko ate sushi.'

The most well-studied aspect marker in Japanese is *te-iru*, whose interpretation is dependent on the lexical aspect of the verb phrase to which it is attached[1]. *Te-iru* cannot combine with stative verbs such as *iru* or *aru* 'be'.[2] With activities and accomplishments, as in (4) and (5), the preferred reading is progressive, which is related to the category of imperfective aspect (Jacobsen 1992; Kindaichi 1950; McClure 1995; Ogihara 1998; Shirai 1998, 2000). However, with achievements, as in (6), *te-iru* denotes a resultative interpretation, which is related to the category of perfective aspect, and a progressive interpretation is ruled out. When achievements are inflected with *te-iru* as in (6), the verb phrase describes the state that obtains ('The plane is at the airport.') as the result of the change of state encoded by the verb. There is also a class of verbs such as *siru* 'come to know' or 'learn' and *niru* 'come to resemble', which are commonly used with *te-iru* (*sit-te-iru* 'know', *ni-te-iru* 'resemble'); an example is shown in (7). While their common *te-iru* forms translate to English stative verbs, the bare Japanese verbs themselves fit into the class of achievements when tests for lexical aspect are applied (see McClure 1995; Shirai 2000 for discussion).

(4) *Taroo ga* *hasit-te-iru.* Activity
Taro NOM run-*te-iru*
'Taro is running.'

(5) *Taroo ga* *hon o* *yon-de-iru.* Accomplishment
Taro NOM book ACC read-*te-iru*
'Taro is reading a book.'

1 Although we focus on the unique interaction between *te-iru* and the lexical aspectual classes of verbs in Japanese, there are also two additional readings of *te-iru*, the experiential and habitual readings, which are available with verbs of any lexical aspectual class (see Fujii 1966; Ogihara 1999; Sugita 2009).

2 An anonymous reviewer points out that some regional dialects of Japanese may allow *te-iru* with the stative verb *iru* 'be.'

(6) *Hikooki ga* *kuukoo ni* *tui-te-iru*. Achievement
plane NOM airport LOC arrive-*te-iru*
'The plane (arrived and) is at the airport.'

(7) *Taroo wa* *omosiroi* *hon o* *sit-te-iru / siru*. Achievement
Taro TOP interesting book ACC know-*te-iru* / know
'Taro knows an interesting book.'/ 'Taro will know of an interesting book.'

The form *te-iru* can itself be inflected for past tense as in the examples in (8) and (9). With accomplishments as in (8), the interpretation of the past form of *te-iru*, *te-ita*, is similar to the English past progressive (*Tarō was reading a book*). With achievements, as in (9), *te-ita* entails completion and the interpretation is most similar to the English past perfective.

(8) *Taroo ga* *hon o* *yon-de-ita*
Taro NOM book ACC read-*te-ita*
'Taro was reading a book.'

(9) *Hikooki ga* *kuukoo ni* *tui-te-ita*.
plane NOM airport LOC arrive-*te-ita*
'The plane had arrived at the airport.'

## 3 Aspect Hypothesis studies in L2 Japanese

The majority of studies that have examined tense and aspect in L2 Japanese have focused on the Aspect Hypothesis (Anderson and Shirai 1994), which predicts that learners are biased in their use of temporal and aspectual forms. We will focus on the two specific predictions listed in (i) and (ii) below (Anderson and Shirai 1996; Bardovi-Harlig and Bergstrom 1996; Shirai 1991) that have been at the center of research on the Aspect Hypothesis in L2 Japanese. We will explain how each prediction has been applied to Japanese.

(i) Learners will first use past/perfective marking with telic verbs such as accomplishments and achievements and will later extend its use to activities and stative verbs.
(ii) For languages that have progressive aspect, learners will first use progressive marking with activity verbs and later extend its use to accomplishments and achievements.

With specific reference to Japanese, the hypothesis in (i) predicts that the past marker –*ta* will be associated with accomplishments and achievements. However, it

is not straightforward to directly apply the claim in (ii) to Japanese because the aspectual marker *te-iru* can encode a progressive meaning with activities and accomplishments but a resultative interpretation with achievements. Shirai (1993, 2002) proposes that because learners often establish one-to-one relationships in linking form to meaning, it is predicted that learners of Japanese will associate *te-iru* with activity verbs (in order to encode progressive meaning) as opposed to with achievements, which are predicted to be strongly associated with past/perfective tense marking. In all L2 Japanese research that has investigated the Aspect Hypothesis, the proposal in (ii) has been taken to mean that *te-iru* will be used first with activity verbs. Extending further from this, Shirai (2002) points out that the prediction in (ii) has also generally been interpreted to mean that the progressive interpretation of *te-iru* (see examples (4) and (5)) will be acquired before the resultative interpretation (see example (6)).

As has been summarized in two previous reviews on the acquisition of aspect in L2 Japanese (see Li and Shirai 2000; Shirai 2002a), several studies have observed these predicted patterns in L2 Japanese learners from a range of L1 backgrounds (primarily English, Chinese, and Korean) and using a range of tasks, including those targeting oral production and others which require the learner to supply the correct morphological form (Koyama 1998, 2004; Kurono 1995; Sheu 1997, 2000; Shibata 1999; Shiokawa 2007; Shirai and Kurono 1998; Sugaya and Shirai 2007). Ishida (2004) is the only study to report a different pattern for *te-iru*. In her longitudinal examination of four learners, she found that learners used *te-iru* more accurately to encode a resultative meaning than to encode a progressive meaning; interestingly, this difference in accuracy held across all test sessions. Ishida proposed that the learners may be more accurate with the resultative use of *te-iru* because it was introduced first in their textbooks while the progressive interpretation was not introduced until four months later. This suggests that both input and instruction impact the learners' ability to use the form accurately.

An interesting question that emerges from this review is what explains the pattern that is generally observed in studies of tense and aspect in L2 Japanese? With the exception of Ishida (2004), why is *te-iru* generally associated with activities and a progressive interpretation and why is the past marker –*ta* associated with telic verb phrases? Anderson and Shirai's (1996) Prototype Hypothesis proposed that learners first acquire the most prototypical members of a category. For past/perfective markers, which encode completion, telic verb phrases, such as accomplishments and achievements are prototypical in that the verb phrases encode endpoints and thus help to define the notion of completion. On the other hand, atelic verb phrases, such as activities, which do not encode an endpoint, are prototypes of temporal and aspectual forms such as the imperfective. The formation of the prototype is argued to be tied to a distributional bias in the input. For example, Shirai (1995) reports that even in native speech, the use of –*ta* is strongly aligned with achievement verbs. Thus, frequency in the input is a possible explanation for the skewed distribution

in learner speech. However, the case of *te-iru* is much trickier because in native speech, *te-iru* is not strongly aligned with activity verbs. Rather, Shirai and Kurono (1998) showed that native speakers use *te-iru* more frequently with achievements to encode a resultative interpretation in learner-directed speech (see also Shirai and Nishi 2005). Thus, the learners' tendency to use *te-iru* with activities cannot be derived from a distributional bias in the input and must be due to other factors.

Recent research has presented evidence that the properties of the L1 may play a role. For example, Sheu (1997) argued that Chinese learners' difficulty with the resultative interpretation of *te-iru* may be tied to their tendency to associate the perfective marker *le* in Chinese with the perfective marker *–ta* in Japanese. Thus, they tend to use and accept Japanese *–ta* in contexts in which Japanese native speakers would use the resultative *te-iru*. Koyama (2004) reports similar findings in which Chinese native speakers had more difficulty than Korean or English native speakers in using *te-iru* with achievement verbs to encode a resultative interpretation.[3] Koyama used a fill-in-the-blank judgment test based on a task used in Kurono (1995) in which learners were asked to choose from a set of inflected forms as in (10) below. The verbs were inflected in the non-past, simple past, *te-iru*, and *te-ita* (past form of *te-iru*) forms. Learners were allowed to choose more than one form. In the example in (10), the 'target form,' which would be selected by native speakers, is in bold.

(10) A: *Doo simasita ka*. ('What's the matter?')
B: *Zitensya no kagi o* __(1)__ *n desu*. ('I actually ____ my bike key.')
A: *Zitensya no kagi? Asoko ni* __(2)__ *yo*. ('Your bike key? It ____ over there.')
B: *E, doko desu ka*. ('What? Where is it?')
A: *Asoko desu. Tukue no sita ni* __(3)__ *yo*. ('It's over there. It __ under the table.')
1. *sagasu, sagasita,* ***sagasiteiru****, sagasiteita* (*sagasu* 'look for')
2. ***arimasu****, arimasita, atteimasu, atteimasita* (*aru* 'be-inanimate')
3. *otimasu, otimasita,* ***otiteimasu****, otiteimasita* (*otiru* 'fall')

Koyama found that the Chinese learners were more likely to choose the simple past *–ta* as opposed to *te-iru* with achievement verbs such as *otiru* 'fall.' In this case, native speakers would use the *te-iru* form (*otiteimasu*) unless the speaker had actually witnessed the key fall under the table, in which case the simple past form

3 Shiokawa (2007) used a similar task to Koyama and also reports a strong association for learners between achievement verbs and the past tense *–ta* even in contexts in which *te-iru* is generally used. What is unique about this study is that the stimuli examines noun-modifying clauses such as *koware-te-iru pasokon* 'broken computer' in which even native speakers will allow the simple past (i.e. *koware-ta pasokon*) in addition to *te-iru*. Unfortunately, the L1 background of the learners is not provided in the article, so we cannot evaluate whether transfer is a potential explanation for the pattern observed in the learners.

(*otimasita*), which the Chinese learners selected, would be used. These subtle distinctions in the use of resultative *te-iru* and the simple past are clearly difficult for L2 learners.

The results of Sheu (1997) and Koyama (2004) show that the resultative use of *te-iru* is difficult in part because of the potential competition from the past form –*ta* in Japanese. Although the association between achievements and –*ta* is predicted by the Aspect Hypothesis, this difficulty is especially salient for Chinese learners who have formed a strong association between the perfective marker *le* in their L1 and the past marker –*ta* in Japanese. This suggests that the properties of the L1 contribute to this association. A similar question can be asked of the association between activities and the progressive use of *te-iru*. Are the high levels of accuracy with the progressive use of *te-iru* related to the fact that most studies have focused on learners whose native language has a progressive form?

A recent study by Sugaya and Shirai (2007) directly addressed this question by comparing the performance of a group of English-speaking learners of Japanese with the performance of a group of Japanese learners whose L1 does not have obligatory progressive marking (German, Russian, Ukrainian, and Bulgarian); Sugaya and Shirai refer to this second group of participants as the 'L1 Non-progressive' group. They predicted an advantage for the progressive interpretation of *te-iru* in comparison to the resultative interpretation of *te-iru* for the English native speakers learning Japanese but did not predict such an advantage for the learners whose L1 does not have an obligatory progressive form. Sugaya and Shirai used a picture description task in which participants were asked to 'spot the difference' between two pictures and a fill-in-the-blank judgment test similar to the one Koyama (2004) used. The judgment task targeted both the progressive and resultative interpretation of *te-iru*; an example targeting the resultative interpretation with the achievement verb *tuku* 'to be attached to' is given in (11) below. Learners were allowed to select more than one option; the target choice (C) is in boldface below. Similar to the example above in (10), a native Japanese speaker would select the *te-iru* form in (C) in the example in (11) unless the speaker had actually observed the moment at which the shirt was stained with lipstick (in which case the simple past form in B would be selected).

| | | |
|---|---|---|
| (11) | Takahashi: | *Are, syatu ni kutibeni ga _____ ne.* |
| | | ('Oh, there's lipstick on your shirt'.) |
| | Yamamoto: | *E, hontoo desuka?* ('Oh, really?') |
| | A. | *tukimasu* (non-past, polite) |
| | B. | *tukimasita* (past, polite) |
| | C. | ***tuiteimasu*** (*te-iru*, polite) |
| | D. | *tuiteimasita* (*te-ita*, polite) |

The learners were divided into two proficiency levels (Higher and Lower) on the basis of their performance on items on the judgment task that targeted the non-past and simple past forms. Importantly, the learners in the L1 English group in each of the two proficiency levels were matched in terms of proficiency to the learners in the 'L1 Non-progressive' group. Thus, any differences which arise between the L1 English group and the 'L1 Non-progressive' group are more likely to stem from differences in the linguistic properties of their native languages, and not simply differences in proficiency level.

The results of the judgment task did not reveal any effects for the learners' L1. In both the L1 English group and the 'L1 Non-progressive' group, the Higher proficiency learners performed well with both the progressive and resultative uses of *te-iru*. The only difference that emerged between the two contexts was that in the sentences targeting the resultative use of *te-iru*, learners often selected the simple past in addition to *te-iru*; for the sentences targeting the progressive use of *te-iru*, learners did not select an alternative. For the lower proficiency groups, there was an advantage for the progressive use of *te-iru* as opposed to the resultative use. However, this advantage emerged for both L1 groups. Thus, the results of the judgment task support the Aspect Hypothesis in that there is an advantage for the progressive use of *te-iru* (activities + *te-iru*) for lower proficiency learners. In addition, there was an overall association between achievements and the simple past for learners at both proficiency levels, and there was not a strong effect of the L1.

However, an L1 effect did emerge in the results of the picture description task, at least for the learners at lower proficiency levels. The learners at a higher level of proficiency in both L1 groups performed well with both the progressive and resultative uses of *te-iru*, with accuracy at about 90%. At lower levels of proficiency, the L1 English learners performed better with the progressive use of *te-iru* (91%) than the resultative use (72%) while the learners in the 'L1 Non-progressive' group struggled with both (progressive: 74%; resultative: 72%). The most salient difference between the two L1 groups was that the learners whose L1 does not have an obligatory progressive form used the simple non-past form in order to encode progressive meaning. This is a clear transfer effect as progressive meaning is encoded by simple present forms in the L1s of the learners tested.

Based on the results of these two tasks, Sugaya and Shirai (2007) propose that L1 transfer is an important factor, but not the only factor, which leads to the advantage for the progressive interpretation of *te-iru* for learners. The results of the production task for the low proficiency learners support a transfer account but recall that on the acceptability judgment task, the advantage for the progressive emerged even for the low proficiency learners in the 'L1 Non-progressive' group. Sugaya and Shirai (2007) propose that the progressive advantage may emerge in part because learners tend to rely on one-to-one mappings between form and meaning (see also Shirai and Kurono 1998). For the progressive meaning in Japanese, *te-iru* is obligatory, but for events which have already occurred, learners may perceive competition between

several grammatical forms, including *te-iru* and the past marker *–ta*. Thus, learners at lower levels of proficiency, who are limited in their form-meaning mappings, will favor the association between *te-iru* and the progressive reading and will show an advantage for the progressive while learners at higher levels of proficiency, who allow more complex relationships between form and meaning, will be able to perform well with both the progressive and resultative meanings. In Japanese, the perceived 'competition' between *te-iru* and *–ta* in the encoding of the resultative meaning may be an important factor in accounting for the patterns that consistently emerge. Thus, these results suggest that L1 transfer indeed impacts L2 acquisition but the specific properties of the target language play an important role as well. Gabriele's studies of Chinese and English-speaking learners of Japanese (Gabriele 2009; Gabriele and McClure 2011), which will be reviewed in the next section, suggest a similar interaction of factors.

# 4 Transfer at "different levels"

## 4.1 Grammatical aspect

Gabriele (2009) takes a somewhat different approach to the examination of transfer in the acquisition of *te-iru*. Similar to the research reviewed above, her study also investigates whether learners can successfully acquire the progressive and resultative interpretations of *te-iru* but in addition, she directly examines whether they can successfully rule out the interpretations that would be allowed by the L1 grammar. The question of whether learners can 'rule out' properties which are allowed in the L1 but are prohibited in the L2 has traditionally been of interest in generative L2 research (e.g. Mazurkewich 1984, Inagaki 2001). With respect to *te-iru*, consider the learning task for an English-speaking learner of Japanese: if the learner transfers from the L1, it is likely that the learner will treat the aspectual marker *te-iru* in Japanese similarly to the progressive marker *be-ing* in English. As both aspectual markers encode progressive aspect in the two languages, transfer from the L1 will facilitate acquisition in L2 Japanese in cases in which *te-iru* encodes a progressive interpretation, specifically with activities and accomplishments. However, consider the sentence in (12), which includes an achievement verb inflected with *te-iru* and encodes a resultative interpretation.

(12) *Hikooki ga* *kuukoo ni* *tui-te-iru.*
plane NOM airport LOC arrive-*te-iru*
'The plane has arrived at the airport.'

In this case, transfer from the L1 will cause difficulty because the learner will incorrectly interpret *te-iru* as progressive and thus, incorrectly interpret the sentence in (12) as *The plane is arriving at the airport* instead of *The plane has arrived at the airport*. In this case, the learner will need to accomplish two different goals in order

to converge on the target interpretation. First, the learner must learn that an achievement verb with *te-iru* encodes a resultative interpretation, and second, they must learn that the interpretation is *exclusively* resultative. In other words, a progressive reading (*The plane is arriving at the airport*), which would be allowed in the L1 English, must be ruled out. Gabriele (2009) designed an interpretation task to target exactly these learning tasks in order to investigate whether it is easier to acquire "new" properties of the L2 than it is to rule out properties that are relevant to the L1 but not the L2.

Participants were given a Story Compatibility task which presented stories in Japanese followed by test sentences. Participants were then asked to determine whether the test sentence was compatible with the story (within a ten second time frame), judging each test sentence on a scale of 1–5 (with "5" representing 'This sentence is completely compatible with this story'). The stories were presented via pictures and audio narration on a computer. An example of a story targeting the achievement verb *tuku* 'arrive' is presented in (13). Similar stories were developed for accomplishment verbs such as *e o kaku* 'draw a picture' as well. For each verb, there were two different story contexts. The two versions began with the same opening narration, but one story depicted an event that was complete as in (13a), and the second story depicted an event that was incomplete, or in progress as in (13b).

(13) *Tuku* 'arrive' (achievement)

Picture 1:

| | | | |
|---|---|---|---|
| *Kore wa* | *tookyoo-yuki no* | *hikooki* | *desu.* |
| this TOP | Tokyo-bound GEN | plane | COP |

| | | |
|---|---|---|
| *Ima* | *yozi* | *desu.* |
| now | four o'clock | COP |

| | | | |
|---|---|---|---|
| *Hikooki wa* | *kuukoo no* | *tikaku* | *desu.* |
| plane TOP | airport GEN | near | COP |

'This is the plane bound for Tokyo. It's 4:00 now. The plane is near the airport.'

a. Complete story context

Picture 2a:

| | | | | |
|---|---|---|---|---|
| *Gozi* | *desu.* | *Zyookyaku wa* | *kuukoo ni* | *imasu.* |
| five o'clock | COP | passenger TOP | airport LOC | be |

'It's 5:00. The passengers are at the airport.'

b. Incomplete story context

Picture 2b:

| | | |
|---|---|---|
| *Kaze ga* | *tuyoi* | *desu.* |
| wind NOM | strong | COP |

| | | | |
|---|---|---|---|
| *Yozi* | *sanzyuppun ni* | *hikooki wa* | *mada* |
| four o'clcock | thirty minutes LOC | plane TOP | still |

| | | |
|---|---|---|
| *sora o* | *hikoo-tyuu* | *desu.* |
| sky ACC | flying-in.the.middle.of | COP |

'The wind is strong. At 4:30 the plane is still in the sky.'

Test sentences

| | | | |
|---|---|---|---|
| **Simple Past**: | *Hikooki wa* | *kuukoo ni* | *tukimasita.* |
| | plane TOP | airport LOC | arrived |
| | 'The plane arrived at the airport.' | | |

Complete Story (ex. 13, Picture 2a): Accept
Incomplete Story (ex. 13, Picture 2b): Reject

| | | | |
|---|---|---|---|
| ***Te-iru***: | *Hikooki wa* | *kuukoo ni* | *tui-te-imasu.* |
| | plane TOP | airport LOC | arrive-*te-iru*.POL |
| | 'The plane is at the airport.' | | |

Complete Story (ex. 13, Picture 2a): Accept
Incomplete Story (ex. 13, Picture 2b): Reject

Learners were then presented with test sentences that targeted either the interpretation of the simple past –*ta* or *te-iru*. With achievement verbs, both the simple past and *te-iru* test sentences should be accepted with the complete contexts and rejected with the incomplete contexts. This is the target pattern predicted for Japanese native speakers. For English-speaking learners, performance on the simple past should be facilitated by the properties of the L1 as the interpretation of the simple past is similar in the two languages. However, for *te-iru*, if English-speaking learners transfer from the L1, they will incorrectly reject the sentence with *te-iru* with the complete context and incorrectly accept the sentence with *te-iru* with the incomplete context. Performance on achievement verbs with *te-iru*, which differ in English and Japanese, was compared to performance on accomplishment verbs, which behave similarly in English and Japanese. With accomplishment verbs, the simple past test sentence should be accepted with the complete contexts and rejected with the incomplete contexts, similar to the achievements. However, the *te-iru* test sentence should show the opposite pattern as it should be accepted with the incomplete context (*painting a portrait*) and rejected with the complete context (*has painted a portrait*) (see Gabriele, 2009 for a complete example). Thus, no difficulty was predicted for the accomplishments for either the simple past or the *te-iru* test sentences as the L1 English can facilitate performance in both cases.

English-speaking learners of Japanese at two proficiency levels, Low (n = 16) and High (n = 17), as well as a group of Japanese native speakers (n = 31) completed the task. Learners were divided into proficiency levels on the basis of scores on selected items from the Japanese Language Proficiency Test (Levels 2 and 3). Learners performed well with the simple past sentences for both accomplishments and achievements, demonstrating the same patterns as the Japanese natives. The results for the accomplishments were also generally in line with the predictions, with learners again showing the same patterns as native speakers. Thus, in all cases where the L1 and L2 are similar, learners at both proficiency levels performed well.

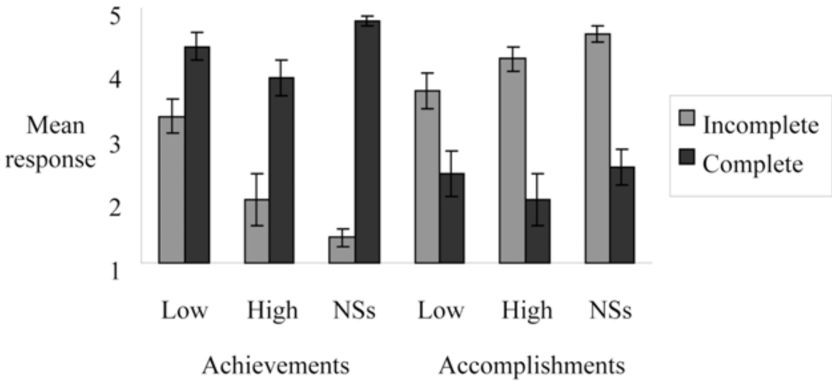



**Figure 1:** L1 English-L2 Japanese: Mean acceptance of accomplishments and achievements with *te-iru* in incomplete and complete contexts (Reprinted from Gabriele, A. 2009. Transfer and transition in the L2 acquisition of aspect. *Studies in Second Language Acquisition* 31. 371–402. Copyright © (2009) Cambridge University Press. Reprinted with permission.)

The results for *te-iru* for both accomplishments and achievements are summarized in Figure 1.

For the *te-iru* sentences with achievements, an interesting pattern emerged. For the complete story context (13a), in which the *te-iru* sentence should be accepted, learners at both proficiency levels performed well. These results suggest that the learners have acquired the resultative interpretation of *te-iru*, even at lower proficiency levels. However, the learners showed more difficulty with the incomplete context (13b), which depicts an event in progress. Incorrect acceptance of the *te-iru* sentences with the incomplete context suggests that the learners can allow a progressive reading for *te-iru*, unlike the Japanese natives. Although there were several high proficiency learners who correctly rejected all *te-iru* sentences with the incomplete context, difficulty was observed among individual learners at both proficiency levels.

The results of Gabriele's (2009) study point to two patterns that are somewhat unique given the results of previous studies. First, in correctly accepting the *te-iru* sentences with the complete context, it is clear that the learners, even those at low levels of proficiency, do not have difficulty with the resultative interpretation of *te-iru*, unlike what has largely been observed in previous studies. It is possible that the reason for this discrepancy is methodological. By presenting learners with a single story context and a single test sentence, the task used in Gabriele (2009) does not put *te-iru* 'in competition' with other grammatical forms such as the simple past. As was reviewed in the previous section, many previous studies used fill-in-the-blank questions in which learners are asked to choose from a selection of inflected verb forms including *te-iru* and the simple past. Given that the learners seem to strongly align completion with the simple past, these tasks may not allow the learners to fully consider *te-iru* as a possibility for encoding the resultative inter-

pretation. Second, Gabriele's results showed that even when learners have acquired the target-like resultative interpretation of *te-iru*, they may nevertheless have difficulty ruling out the progressive interpretation that is allowed by the L1 grammar. This provides a new perspective with respect to the factors that may cause difficulty in the acquisition of achievements with *te-iru*.

Although the results of Gabriele (2009) suggest a role for transfer, the results of a subsequent study using the same experimental paradigm suggested that the specific properties of the target language may also play a very important role in determining relative ease or difficulty of acquisition. In a follow-up study, Gabriele and McClure (2011) tested Chinese learners of Japanese using the same Story Compatibility task that was described above (see also Gabriele 2008). The Chinese learners (n = 46) were also tested using the same proficiency test as the English-speaking learners and all were classified as advanced learners. All learners were tested in Japan and used Japanese on a daily basis. The results of the Story Compatibility task for the sentence types described above showed native-like performance on all categories. Chinese learners performed well with the simple past with both accomplishments and achievements, correctly accepting simple past sentences with complete contexts and rejecting them with incomplete contexts. As for *te-iru*, Chinese learners showed that they had acquired the target-like interaction between lexical aspect and grammatical aspect. For achievements with *te-iru*, they correctly accepted the *te-iru* sentences with the complete context and correctly rejected them with the incomplete context; they also correctly showed the opposite pattern with accomplishments.

However, there was one case in which Chinese learners showed difficulty. Gabriele and McClure (2011) tested the same story contexts but also included test sentences targeting *te-ita*, the past form of *te-iru*. This is the most interesting context in which to examine Chinese learners of Japanese as Chinese does not have a grammatical marker of tense, and thus the examination of the interpretation of *te-ita* addresses the question of whether Chinese learners can successfully acquire a linguistic property that is not instantiated in the L1.

*Te-ita* sentences for both achievements and accomplishments are presented below in (14) and (15). Just as in Gabriele (2009), participants were asked to judge each test sentence on a scale of 1–5, evaluating whether or not the sentence was compatible with the story they had just listened to (with "5" representing 'This sentence is completely compatible with this story'). The full story contexts for the achievement verb 'arrive' are given in (13) (see Gabriele, 2009 and Gabriele and McClure, 2011 for the story contexts for the accomplishment 'paint a picture').

(14) *Hikooki wa* *kuukoo ni* *tui-te-imasita.*
plane TOP airport LOC arrive-*te-ita*.POL
'The plane had arrived at the airport.'

<u>Complete Story (ex. 13, Picture 2a)</u>: Accept
<u>Incomplete Story (ex. 13, Picture 2b)</u>: Reject

(15) *Ken wa* *kazoku no* *e o* *kai-te-imasita.*
Ken TOP family GEN picture ACC draw-*te-ita*.POL
'Ken was painting a picture of his family.'

Complete Story: Accept
Incomplete Story: Accept

In (14), with the achievement verb *tuku* 'arrive', the *te-ita* sentence is interpreted as a past resultative (*The plane had arrived at the airport*) and is only compatible with the complete context. As is summarized in Figure 2, Chinese learners performed well on sentences such as (14), correctly accepting *te-ita* with achievements with the complete context and rejecting them with the incomplete context.

However, as is also summarized in Figure 2, Chinese learners showed the same pattern with the *te-ita* sentences with accomplishments as in (15), although that is not the target-like response. Accomplishments with *te-iru* encode a progressive reading and thus the sentence in (15) is interpreted as *Ken was painting a portrait*. The sentence in (15) should be accepted with *both* the incomplete and complete contexts as the past *te-ita* sentences with accomplishments can refer to events that were ongoing in the past, regardless of whether or not they were completed. Thus, successful interpretation of (15) requires that the learner can tease apart the relative contribution of (past) tense and (perfective) aspect. Native speakers of Japanese indeed accepted sentences such as (15) with both the incomplete and complete contexts but the Chinese learners accepted sentences such as (15) largely with the complete context only. The results for *te-ita* are interesting given that the learners were successful with the present tense form of *te-iru* with both verb classes.

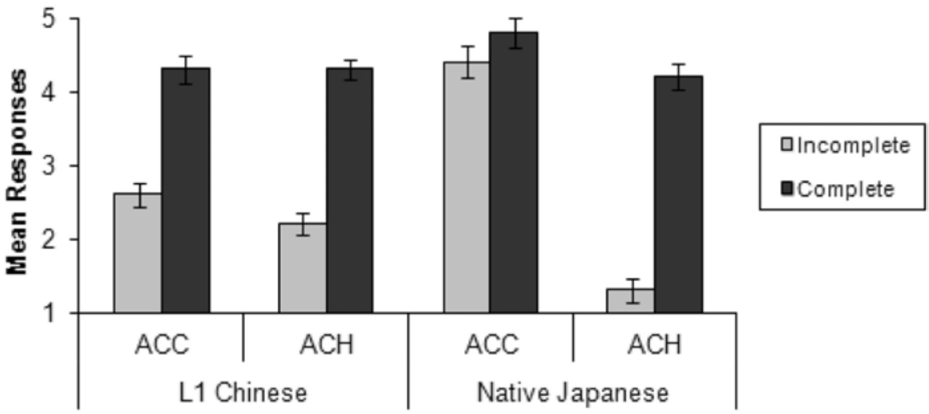



**Figure 2:** L1 Chinese-L2 Japanese: Mean acceptance of accomplishments and achievements with *te-ita* in incomplete and complete contexts (Reprinted from Gabriele, Alison. 2008. Mapping between form and meaning: A case of imperfect L2 acquisition. IEICE Technical Report 108(184). 101–106. Copyright © 2008 IEICE. This figure also appeared in Gabriele and McClure 2011)

In trying to account for these results, Gabriele and McClure first consider the possibility that Chinese learners confounded past tense and perfective aspect, exclusively allowing *te-ita* to refer to completed events regardless of the verb phrase (accomplishment or achievement). Because Chinese does not grammatically encode tense, but does encode perfective aspect, this account would make sense if the learners simply cannot extend beyond the aspectual resources of the L1. This would suggest that features not instantiated in the learners' L1 cannot be acquired to native-like levels. However, Gabriele and McClure rule out this account for two reasons. First, they present results from Gabriele and Maekawa (2008) that show that Chinese learners of English can successfully interpret accomplishments in the past progressive (*Ken was painting a portrait of his family*), suggesting that Chinese learners do not uniformly interpret the past tense as perfective. Rather, the difficulty is restricted to Japanese *te-ita*. Second, they present results from Gabriele (2005), which show that L1 English learners of Japanese also have difficulty with accomplishments with *te-ita*. These results are summarized in Figure 3.

The results for the L1 English learners are surprising in that the interpretation of accomplishments with the English past progressive and Japanese *te-ita* is similar, at least with respect to the contexts tested in this task; nevertheless, English-speaking learners and Chinese-speaking learners show the same pattern of results in L2 Japanese. Gabriele and McClure propose that the morphological encoding of tense and aspect in the *te-ita* form may influence the ease with which those two semantic concepts can be teased apart, regardless of the properties of the learners' L1. In Japanese, the forms *te-iru* and *te-ita* encode both tense and aspect and learners

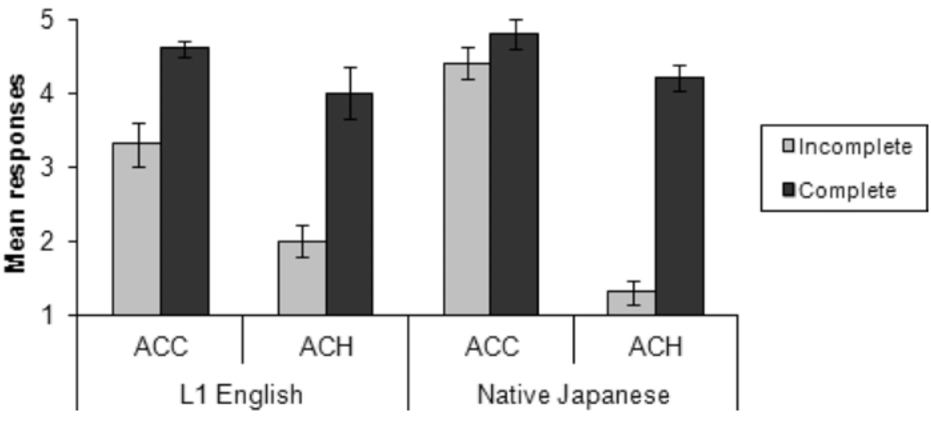



**Figure 3:** L1 English-L2 Japanese: Mean responses to accomplishments and achievements with *te-ita* (Reprinted from Gabriele, A. and McClure, W. 2011.Why some imperfective are interpreted imperfectly: A study of Chinese learners of Japanese. *Language Acquisition* 18. 39–83. Reprinted by permission of the publisher, Taylor & Francis Ltd, http://www.tandf.co.uk/journals)

may struggle to tease apart each semantic concept in order to derive for *te-ita* (at least for accomplishments) the interpretation of 'past ongoing'. In contrast, for the English past progressive, an auxiliary verb independently carries the features of tense and agreement. In this case, separate morphemes encode tense and aspect (*is/was* paint*ing* a portrait) and the main verb itself in inflected only for progressive aspect. Thus, it is possible that the mapping between form and meaning is facilitated in English by this one-to-one correspondence.

In summary, the recent research on transfer with respect to grammatical aspect suggests that the L1 may be one important factor to consider in the L2 acquisition of aspect but it is certainly not the only determinant of ease or difficulty of acquisition. First, the specific learnability scenario is important to consider. Gabriele's (2009) results suggest that it is easier to acquire new properties in the L2 than it is to rule out properties that are instantiated in the L1 but not the L2. In addition, both Sugaya and Shirai (2007) and Gabriele and McClure (2011) suggest that the specific linguistic properties of the target L2 form in question may play an equally important role in determining ease of acquisition. It is likely that more transparent mappings between form and meaning in the morphological encoding of tense and aspect will facilitate acquisition. Note that these generalizations are only possible due to the comparisons between multiple (proficiency controlled) L1 groups in both Sugaya and Shirai's (2007) and Gabriele and McClure's (2011) studies.

## 4.2 Lexical aspect

While the research summarized above focused primarily on transfer at the level of grammatical aspect, Nishi (2008) examines the extent to which crosslinguistic differences at the level of lexical aspect may play a role in L2 acquisition. This is an important question as a comprehensive understanding of L1 influence requires an understanding of what specific linguistic properties are candidates for transfer. As we reviewed in the introduction, although the semantic features underlying the four lexical aspectual classes are arguably universal, the specific classification of a given verb may differ depending on the language. In her dissertation, Nishi examines the acquisition of *te-iru* in Japanese by native speakers of English, Chinese, and Korean at three different levels of Japanese proficiency. We focus here just on the results for the English-speaking learners. Learners were tested using a paper and pencil translation task and a production task in which they were asked to describe pictures. The translation task, which we will focus on here, examines the progressive and resultative interpretations of *te-iru*, similar to previous studies, but with respect to the resultative interpretation, she compares verbs that are either similar or different in terms of their lexical aspectual classification in the L1 and L2.

Selected examples from a subset of the categories on the translation test for English-speaking learners of Japanese are provided in (16)–(17) below. The Japanese

sentence that the learners were provided with is presented below in italics and the English sentence that the participants were asked to compare it to is underlined. (The meaning of the Japanese test sentence is also provided below in quotations but was not provided to the learners.)

Participants were asked to decide if the Japanese sentence and the English sentence encoded the same meaning. The test items in (16)–(17) all target the resultative interpretation of *te-iru* but differ with respect to whether there is a discrepancy in lexical aspect between Japanese and English. In (16), there is no discrepancy as the verbs tested are achievements in both languages. The sentence type in (16a) targets whether learners have learned that achievements + *te-iru* encode a resultative interpretation and should be accepted as a good match with the English sentence. In contrast, the sentence in (16b) presents a 'mismatch' because the English sentence describes a resultant state while the Japanese sentence is in the simple present, which encodes a futurate interpretation. (16c) is also a mismatch between the Japanese and English sentences because achievement verbs + *te-iru* do not allow a progressive interpretation.

(16) a. *kuru* 'come'/*come* (achievements, no discrepancy): accept
*Matuda san wa kotira ni ki-te-imasu.*
<u>Ms. Matsuda is here (as a result of coming)</u>.
'Matsuda is here (as a result of coming).'

b. *sinu* 'die'/*die* (achievements, no discrepancy): reject
*Ano tori wa sinimasu.*
<u>That bird is dead (as a result of dying)</u>.
'That bird is going to die.'

c. *sinu* 'die'/*die* (achievements, no discrepancy): reject
*Ano kanzya wa sin-de-imasu.*
<u>That patient is dying</u>.
'That patient is dead.'

In contrast, the sentences in (17) present a discrepancy in lexical and grammatical aspect. The verbs included are achievements in Japanese but are activities in English. In terms of grammatical aspect, with verbs such as *noru* 'ride' or *suwaru* 'sit', *te-iru* encodes a resultant state (state of sitting, being on a bus) while the progressive in English encodes an ongoing activity. However, in (17a), the Japanese sentence should be accepted as a good match to the English sentence as there is an overlap in the situations that can be described by the two sentences. In contrast, the sentence pair in (17b) should be rejected as a mismatch because the Japanese simple present sentence encodes a futurate reading while the English sentence describes an activity ongoing in the present.

(17) a. *noru* 'ride' (achievement)/*ride* (activity): accept
*Yamada san wa basu ni not-te-imasu.*
Ms. Yamada is riding a bus (= Ms. Yamada is on a bus).
'Yamada is riding a bus.'

b. *tatu* 'stand' (achievement)/*stand* (activity): reject
*Honda san wa asoko ni tatimasu.*
Honda is standing over there.
'Honda is going to stand over there.'

The results suggest that differences between the L1 and L2 in grammatical aspect play a stronger role than differences in lexical aspect. Despite the fact that verbs in the sentences in (16) are achievements in both languages, the test items were still very difficult for low proficiency learners, who performed below chance, and intermediate proficiency learners, whose performance was just above chance. This difficulty suggests it is the resultative interpretation of *te-iru* that presents a strong challenge to the learners, showing the influence of L1–L2 differences at the level of grammatical aspect. However, the advanced L1 English group performed well in this category, showing that the resultative interpretation can be mastered at higher levels of proficiency.

In contrast, all of the learners performed better with the sentence pairing in (17a), despite the discrepancy between the L1 and L2 with regard to lexical aspect. In this case, target-like performance is facilitated by the fact that there is an overlap between Japanese and English in the situations that can be described with resultative *te-iru* and progressive *be+ing*, particularly with verbs such as *sit* and *ride*. In contrast, the results for test items such as (17b) looked similar to the results observed for the items in (16) with low and intermediate proficiency learners showing difficulty but advanced learners performing very well. Similar to the sentences in (16), it is likely that differences in grammatical aspect are responsible for the difficulty at lower levels of proficiency. If lower proficiency learners incorrectly interpret the Japanese sentence in (17b) as *Honda stands over there*, they may incorrectly accept it as a match for the English sentence *Honda is standing over there* as there would be a perceived overlap in the situations that can be described with resultative *te-iru* and progressive *be+ing*. It is not until advanced levels of proficiency that the learners correctly interpret the Japanese sentence in (17b) as futurate (*Honda is going to stand/will stand over there*) and thus reject it as a mismatch for the English sentence.

Overall, these results suggest that it is difficult to tease apart effects due solely to differences between the L1 and L2 at the level of lexical aspect when there are also differences at the level of grammatical aspect. But the fact that there is still evidence of difficulty when lexical aspect between the L1 and L2 is held constant, e.g., (16), and relatively less difficulty in a case of a lexical aspect mismatch (17a) suggests

that differences at the level of grammatical aspect play a more important role with respect to transfer.[4]

## 4.3 Defining an event in L2 Japanese

The studies reviewed above all focus on the interaction between lexical and grammatical aspect, in order to understand whether learners can derive the correct interpretation of an aspectual marker such as *te-iru* depending on the lexical semantics of the verb phrase it is attached to. This question however assumes that learners can correctly classify the lexical aspect of the verb phrase itself, which is an issue that should not be taken for granted. In a language such as English, a noun phase such as *wrote the letter* is unambiguously telic (an accomplishment) and a noun phrase such as *wrote letters* is unambiguously atelic (an activity). Crucially, this difference in telicity is encoded in the morphosyntactic form of the direct object (*wrote the letter* vs. *wrote letters*). The verb phrase *wrote the letter* is telic because the direct object *the letter* specifies a specific quantity, and thus, an endpoint (Dowty 1991; Krifka 1992; Tenny 1994). In contrast, an atelic verb phrase, such as *wrote letters*, does not define an endpoint. It is appropriate to describe any subinterval of ‘writing letters’ (regardless of whether or not there is ever a complete letter) with the verb phrase *wrote letters*.

If we consider the facts in Japanese on the other hand, an interesting contrast arises. Unlike English, Japanese does not have plural morphology or determiners, and bare nouns such as *tegami* ‘letter’ can appear freely in argument positions as in (18). In English, there is a contrast between ‘count’ nouns such as *letter*, which can never appear in bare form (**John wrote letter*), and ‘mass’ nouns such as *juice* which can (*John drank juice*). By contrasting (18) and (19), it is clear that Japanese does not morphosyntactically encode a contrast between mass and count nouns. Importantly, bare nouns in Japanese are underspecified with respect to number. Thus, a verb phrase such as *tegami o kakimasita* ‘wrote letter’ in (18) can be interpreted as either telic ‘wrote the letter’ or atelic ‘wrote letters’ depending on the context. In order to explicitly encode number, a classifier must be used, as in (20).

**4** It is also important to point out that the difficulty presented by the L1–L2 differences in grammatical aspect may have been exacerbated by the methodology used in which learners were asked to compare directly between the L1 and L2, thus encouraging the learners to interpret the L2 via the L1. This may be particularly true of written tasks in which learners can directly compare the morphological forms in the two languages. Despite these methodological issues, the results observed in Nishi’s study, particularly for the resultative interpretation of *te-iru* with achievements, are generally in line with the results of other studies such as Gabriele (2009), which showed difficulty with the resultative interpretation at lower levels of proficiency and more target-like performance at advanced levels.

| | | | | |
|---|---|---|---|---|
| (18) | *Samu wa* | *tegami o* | *kakimasita.* | Count |
| | Sam TOP | letter ACC | wrote | |

‘Sam wrote letter.’
‘Sam wrote a/the/some letter(s).’

| | | | | |
|---|---|---|---|---|
| (19) | *Samu wa* | *zyuusu o* | *nomimasita.* | Mass |
| | Sam TOP | juice ACC | drank | |

‘Sam drank juice.’

| | | | | | |
|---|---|---|---|---|---|
| (20) | *Samu wa* | *san-bai no* | *zyuusu o* | *nomimasita.* | Mass+Classifier |
| | Sam TOP | three-CLF GEN | juice ACC | drank | |

‘Sam drank three glasses of juice.’
(examples from Gabriele 2010)

Gabriele (2010) examined whether English native speakers can acquire the target-like interpretation of a verb phrase with a bare count noun such as (18) given the lack of morphosyntactic cues in Japanese and the learners’ reliance on these cues for interpretation of telicity in the L1. Very little L2 research has examined learning scenarios in which the L2 learner must move from a L1 in which a meaning is encoded explicitly by a morpheme to a L2 in which the same meaning needs to be derived from the context. This is an important issue in that it allows us to examine the specific conditions under which it is and is not possible to overcome transfer.

Gabriele (2010, Study 1) presented both intermediate (n = 38) and advanced (n = 7) learners of Japanese with an interpretation task that targeted bare nouns. Sentences with classifiers, which will not be discussed here, were also included as a control. Participants looked at pictures and listened to short stories narrated in Japanese. Two versions of each story were presented: a version in which the event came to completion (complete) and a version in which the event was terminated (incomplete). Following each story, participants were asked to judge each test sentence on a scale of 1–5, evaluating whether or not the sentence was compatible with the story they had just listened to (with “5” representing ‘This sentence is completely compatible with this story’).

The task included two types of bare nouns: ‘count’ nouns such as *kaado* ‘card’ (21) that are obligatorily specified for number in English and ‘mass’ nouns such as *zyuusu* ‘juice’ (22) that can appear ‘bare’ in both English and Japanese and are interpreted similarly. Examples of these test items are given below (21)–(22) with predictions for the Japanese native speakers given for each test sentence. In this experiment, all of the stories involved multiple objects (e.g. four cards) and in all of the incomplete contexts, at least one of the objects was clearly complete. With this type of context, the bare noun is unambiguously acceptable for Japanese natives because the story involves at least one complete object (see Gabriele (2010) for further discussion of this point). Thus, the experiment tests whether Japanese learners, like native speakers, can accept a verb phrase such *kaado o kakimasita*

'wrote card' in a context in which not all cards have been written or whether they interpret the verb phrase to refer maximally to all of the cards in the context (*wrote the cards*), in which case a sentence such as (21a) would be rejected with the incomplete context.

(21) Count Noun

Picture 1–2: *Today is Ken's birthday. He received four presents. He wants to write thank you cards to his friends. Ken writes three cards. Then he starts to write the last card.*

Picture 3a, Complete Story Context: *He finishes the last card. Then he gives the cards to his friends.*

Picture 3b, Incomplete Story Context: *But Ken has to go to school. He cannot finish the fourth card.*

| | | | |
|---|---|---|---|
| *Ken wa* | *tanzyoobi ni* | *<u>kaado</u> o* | *kakimasita.* |
| Ken TOP | birthday LOC | card ACC | wrote |

'Ken wrote card on his birthday.'

<u>Complete Story (ex. 21, Picture 3a)</u>: Accept
<u>Incomplete Story (ex. 21, Picture 3b)</u>: Accept

(22) Mass Noun

Picture 1–2: *John drinks a lot. After school he pours three glasses of juice. He drinks two glasses of juice. Then he starts to drink the third glass.*

Picture 3a, Complete Story Context: *He finishes the third glass of juice. Then he puts the empty glasses in the sink.*

Picture 3b, Incomplete Story Context: *He cannot finish the third glass. He pours the rest of the juice in the sink.*

| | | | |
|---|---|---|---|
| *Zyon wa* | *gakkoo no* | *ato <u>zyuusu</u> o* | *nomimasita.* |
| John TOP | school GEN | after juice ACC | drank |

'John drank juice after school.'

<u>Complete Story (ex. 22, Picture 3a)</u>: Accept
<u>Incomplete Story (ex. 22, Picture 3b)</u>: Accept

The results for the sentence types above are summarized in Figure 4.

The Japanese native speakers performed as predicted, accepting sentences (with scores near 5) with both "count" and "mass" bare nouns regardless of whether the context was complete or incomplete. The Japanese learners performed well with the "mass" nouns, which are interpreted similarly in English and Japanese, but have more difficulty with the "count" nouns. An analysis of the individual results showed that while there were some intermediate and advanced learners who correctly accepted the bare "count" nouns in (21) with the incomplete story context, many

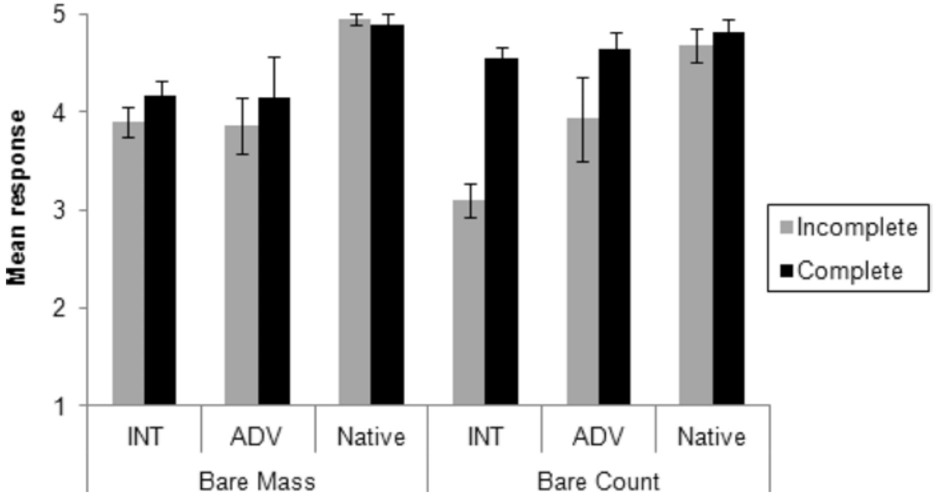



**Figure 4:** L1 English L2 Japanese: mean responses for sentences targeting bare nouns with incomplete and complete contexts (Reprinted from Gabriele, A. 2010. Deriving meaning through context: The interpretation of bare nominals in L2 Japanese. *Second Language Research* 26. 379–405. Reprinted with permission from Sage Publishing Ltd.)

incorrectly rejected these sentences. These learners appear to have difficulty allowing the bare noun to refer to a subset of the objects mentioned in the story. Gabriele (2010) proposed that this interpretation may arise if, in the absence of morphosyntactic cues which can disambiguate the meaning, the learners take the bare noun to refer maximally (*wrote the cards*) to all of the objects in the context.

With respect to the L2 acquisition of tense and aspect, these results show that the interpretation of events in a language like Japanese may be influenced by crosslinguistic differences at a more refined level than has been considered previously. The syntax and semantics of the nominal system may also play an important role. As there is very little research on the question of how learners derive meaning through context, as opposed to through morphosyntactic and syntactic cues, a language such as Japanese provides a promising testing ground for further research on L1 influence in this area (e.g. Ananth 2007).

## 5 L2 processing of tense and aspect

All of the studies reviewed above have used offline behavioral measures and have focused on learners' production and comprehension of tense and aspect morphology. These studies have not examined how the processing of tense and aspect unfolds in real time, exploring the extent to which learners can use temporal and aspectual

information in the course of online sentence processing and whether L2 processing is similar to that of native speakers. The processing of tense and aspect is a new domain of research, even in the monolingual sentence processing literature, but holds promise for the examination of L2 learners as well. A recent pair of studies by Long, Ono, and Sakai (2010) and Long, Nakaishi, Ono, and Sakai (2012) examines exactly these questions.

Long et al. (2010) used a self-paced reading task in which participants read sentences one word at a time on a computer screen; reading times were measured for each word of the sentence. Longer reading times in specific experimental conditions on self-paced reading tasks are taken to indicate more effortful processing. Examples of the stimuli from one of the experiments is presented in Table 2. This experiment examined whether native speakers are sensitive to the compatibility between an adverb and either a telic or an atelic verb in online processing. Durational adverbs such as *zyuppunkan* 'for ten minutes' are usually compatible with atelic verbs (see A) while limiting adverbs such as *zyuppun de* 'in ten minutes' are usually compatible with telic verbs (see D). Thus, Long et al. (2010) predicted faster reading times at the verb for the 'compatible' conditions in A and D than in the conditions in B and C, which pair limiting adverbs such as *zyuppun de* 'in ten minutes' with atelic verbs (B) and durational adverbs such as *zyuppunkan* 'for ten minutes' with telic verbs (C).

The results were in line with these predictions. Reading times at the verb in region 4 revealed a reading time slowdown in condition B as compared to A and a reading time slowdown in condition C as compared to D, suggesting that aspectual incompatibility leads to an increased processing burden. These results suggest that Japanese native speakers process aspectual information incrementally, using the aspectual information from the adverbs to make predictions about the grammatical properties of elements downstream in the sentence, such as the telicity of the verbs. Their follow-up study examined whether L2 processing would show the same sensitivity to aspectual information in online processing.

Long et al. (2012) conducted a similar experiment with Chinese learners of Japanese (n = 24), focusing on learners at very high proficiency levels who had been living in Japan for an average of 6 years and who had passed the highest level of the Japanese Language Proficiency Test (Level 1). The results for the learners showed a different pattern from the native speakers. The L2 learner data showed evidence of a reading time slowdown, in the region following the verb in Region 5. In self-paced reading tasks, it is not uncommon to see effects emerge in this 'spillover' region, one region following the critical region (see Mackey and Gass 2012). However, it is the pattern of the slowdown that differs qualitatively for the two groups. For the learners, there was a main effect of adverbial type, which was the result of longer reading times for the conditions with the durational adverbs such as *zyuppunkan* 'for ten minutes' (A and C) as opposed to the conditions with the limiting adverbs such as *zyuppun de* 'in ten minutes' (B and D). Thus, at least

**Table 2:** Long et al. (2010, 2012): Stimuli from Experiment 1

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| A. Durational | ***zyuppunkan*** | *sensyu ga* | *taiikukan de* | ***rensyuusita*** *to* | *zyosyu ga* | *kantoku ni* | *kyootyoosita* |
| Adverb & Atelic Verb | [for ten minutes] | [athlete NOM] | [gym LOC] | [practiced COMP] | [assistant NOM] | [coach DAT] | [emphasized] |
| B. Limiting | ***zyuppun de*** | *sensyu ga* | *taiikukan de* | ***rensyuusita*** *to* | *zyosyu ga* | *kantoku ni* | *kyootyoosita* |
| Adverb & Atelic Verb | [in ten minutes] | [athlete NOM] | [gym LOC] | [practiced COMP] | [assistant NOM] | [coach DAT] | [emphasized] |
| C. Durational | ***zyuppunkan*** | *sensyu ga* | *kyuuzyoo o* | ***soozisita*** *to* | *zyosyu ga* | *kantoku ni* | *kyootyoosita* |
| Adverb & Telic Verb | [for ten minutes] | [athlete NOM] | [stadium ACC] | [cleaned COMP] | [assistant NOM] | [coach DAT] | [emphasized] |
| D. Limiting | ***zyuppun de*** | *sensyu ga* | *kyuuzyoo o* | ***soozisita*** *to* | *zyosyu ga* | *kantoku ni* | *kyootyoosita* |
| Adverb & Telic Verb | [in ten minutes] | [athlete NOM] | [stadium ACC] | [cleaned COMP] | [assistant NOM] | [coach DAT] | [emphasized] |

in this online measure, the learners show sensitivity to the type of adverb in the sentence (durational versus limiting) but do not show sensitivity to the telicity of the verb or to the compatibility between a specific type of adverb and the telicity of the verb.[5] Long et al. propose that the learners may have found the limiting adverbs such as *zyuppun de* 'in ten minutes' (B and D) to be more compatible with the past marker –*ta*, which is present across conditions, than the durational adverbs in A and C. They suggest that learners may be more affected by grammatical aspect (perfective/past marking) than lexical aspect (telicity) perhaps due to the prominence of grammatical aspect in the L1 Chinese. Although the interpretation of these results is complex, they show that there are interesting reasons to continue to examine learners' online performance in this domain, particularly given the very high proficiency level of the learners. In this way, work on L2 tense and aspect can contribute to the investigation of broader research questions such as the extent to which L2 processing can ever mirror that of native speakers (e.g. Clahsen and Felser 2006; Sawasaki and Kashiwagi-Wood's chapter in this volume).

## 6 Conclusion and future directions

The most recent research conducted on tense and aspect in L2 acquisition has focused in large part on the extent to which the properties of the learner's L1 influences L2 acquisition, both to further examine the patterns frequently observed in research on the Aspect Hypothesis (Sugaya and Shirai 2008) and to better understand the process of transfer itself: Under what learnability conditions can transfer be overcome (Gabriele, 2009, 2010)? Is it easier to acquire "new" properties of the L2 than it is to rule out properties that are relevant to the L1 but not the L2 (Gabriele 2005, 2009; Nishi 2008)? Can grammatical categories such as tense, which are not instantiated in the learner's L1, be acquired successfully in the L2 (Gabriele and McClure 2011)? Is transfer more prevalent at the level of lexical aspect than at the level of grammatical aspect (Nishi 2008)? Importantly, several of these studies have addressed these questions by comparing learners from different L1 backgrounds (Koyama 2004; Nishi 2008; Gabriele and McClure 2011; Sugaya and Shirai 2007) who have crucially been controlled for proficiency level. This body of research shows a united interest in transfer for L2 researchers working in both functional and generative frameworks and also demonstrates that work on L2 Japanese has been "leading the way" for examinations of transfer in the domain of tense and aspect.

Although the number of studies is still relatively small, several generalizations emerge. First, although there is evidence for transfer at the level of the noun phrase

**5** The results of an additional interpretation task in Long et al. (2012) suggest that the learners are sensitive to verbal telicity in offline tasks.

(Gabriele 2010) and at the level of grammatical aspect (Gabriele 2005, 2009; Nishi 2008; Sugaya and Shirai 2007), there is clearly evidence that L1–L2 differences can be overcome at high levels of proficiency (Gabriele 2009; Gabriele and McClure 2011) and that transfer may be more prevalent in contexts in which learners need to "rule out" properties of the L1 that are not instantiated in the L2 (Gabriele 2005, 2009; Nishi 2008). Nevertheless, a very recent study on the processing of tense and aspect suggests that learners may be limited in the type of aspectual information they can use in the course of online processing and that these limitations may be influenced by the properties of the native language (Long et al. 2012). Future research should continue to compare learners' performance on offline and online tasks in order to further examine how task demands impact learners' ability to access temporal and aspectual information.

Second, L1 transfer is not the only factor that determines relative ease or difficulty of acquisition. Both Sugaya and Shirai (2007) and Gabriele and McClure (2011) present cases in which learners whose L1s differ show similar patterns in L2 acquisition, arguably due to the specific properties of the target language. Although work on child language acquisition in this domain is limited, it would be interesting to compare L1 and L2 learners directly on these linguistic properties (e.g., *te-ita*) in order to see whether L2 learners are following similar paths of development as L1 learners in the acquisition of these properties.

Another area that would be interesting to investigate in the domain of L2 tense and aspect is the interaction of markers of grammatical aspect with adverbs. For example, *te-iru* allows a habitual reading with verbs of any lexical aspectual class if an adverb such as *maitosi* 'every year' is present as is shown in (23) (Sugita 2009).

| | | | | |
|---|---|---|---|---|
| (23) | *Mari wa* | *maitosi* | *igirisu ni* | *it-te-i-ru.* |
| | Mari TOP | every year | England LOC | go-*te-iru* |
| | 'Mari goes to England every year.' | | | |

Very few studies of tense and aspect in L2 Japanese have examined these extended meanings (see Sheu 2000; Shirai 2002b) but it is a very interesting domain to investigate, particularly with respect to transfer, as recent studies of L2 English (Gabriele and Canales 2011) have suggested that these extended meanings may not transfer from the L1. Thus, this work would shed light on what linguistic properties transfer in L2 acquisition and what hypotheses learners formulate with respect to the properties of the L2.

In summary, recent work on L2 Japanese in the domain of tense and aspect has fortunately moved the field forward in our understanding of the factors that influence both transfer and ultimate attainment in L2 acquisition. Yet the range of structures and range of methods that have been used to examine these questions is still quite limited. Future research should continue to investigate some of the lesser studied domains within the grammar such as the properties of the noun phrase and

the role of temporal adverbials. In addition, researchers should explore methodologies such as self-paced reading and eye-tracking, which will allow us to see how learners compose an interpretation online. An extension of this line of research beyond the traditional properties that have been investigated and into the domain of psycholinguistics will shed more light on the possibilities and limitations of L2 acquisition and processing.

## Acknowledgments

We are grateful to José Alemán Bañón, Natalie Pak, Keith Swafford, and Luyan Wang for their help in preparing this chapter for publication. We are also grateful to William McClure and Gita Martohardjono who provided feedback on many of the papers and research projects described in this chapter. Several of the projects described in this chapter were supported by grants to the first author from the National Science Foundation (BCS:#0345697) and the General Research Fund (Grant #2301165 and #2301338) at the University of Kansas. We thank Mineharu Nakayama and two anonymous reviewers for their helpful feedback which greatly improved this chapter. All remaining errors are our own.

## References

Ananth, Priya. 2007. *Acquisition of tense and aspect in* Toki *'when' clauses in Japanese as a second/foreign language.* Columbus, OH: Ohio State University dissertation.

Andersen, Roger W. and Yasuhiro Shirai. 1994. Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition* 16. 133–156.

Anderson, Roger W. and Yasuhiro Shirai. 1996. Primacy of aspect in first and second language acquisition: The pidgin/creole connection. In William C. Ritchie and Tej K. Bhatia (eds.), *Handbook of second language acquisition*, 527–570. San Diego, CA: Academic Press.

Bardovi-Harlig, Kathleen and Anna Bergström. 1996. The acquisition of tense and aspect in SLA and FLL: A study of learner narratives in English (SL) and French (FL). *Canadian Modern Language Review* 52. 308–330.

Bohnemeyer, Jurgen and Mary Swift. 2004. Event realization and default aspect. *Linguistics and Philosophy* 27(3). 263–296.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems.* Cambridge: Cambridge University Press.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67. 547–619.

Fintell, Kai von and Lisa Matthewson. 2008. Universals in semantics. *The Linguistic Review* 25. 139–201.

Fujii, Tadashi. 1966. Doushi + te iru no imi [The Meaning of verb + te iru]. *Kokugo Kenkyuu Shitsu* 5. Reprinted In Haruhiko Kindaichi (ed.), 1976. Nihongo dōshi no asupekuto [The aspect of Japanese verbs], 97–116. Tokyo: Mugi Shobo.

Gabriele, Alison. 2005. *The acquisition of aspect in a second language: A bidirectional study of learners of English and Japanese*. New York: Graduate Center, City University of New York dissertation.

Gabriele, Alison. 2008. Mapping between form and meaning: A case of imperfect L2 acquisition. *IEICE Technical Report* 108(184). 101–106.
Gabriele, Alison. 2009. Transfer and transition in the L2 acquisition aspect. *Studies in Second Language Acquisition* 31. 371–402.
Gabriele, Alison. 2010. Deriving meaning through context: The interpretation of bare nominals in L2 Japanese. *Second Language Research* 26. 379–405.
Gabriele, Alison and Alonso Canales. 2011. No time like the 'present': Examining transfer at the interfaces in second language acquisition. *Lingua* 121. 670–687.
Gabriele, Alison, Gita Martohardjono and William McClure. 2003. Why *swimming* is just as difficult as *dying* for Japanese learners of English. *ZAS Papers in Linguistics* 29. 85–104.
Gabriele, Alison and William McClure. 2011. Why some imperfective are interpreted imperfectly: A study of Chinese learners of Japanese. *Language Acquisition* 18. 39–83.
Inagaki, Shunji. 2001. Motion verbs with goal PPs in the L2 acquisition of English and Japanese. *Studies in Second Language Acquisition* 2. 153–170.
Ishida, Midori. 2004. Effects of recasts on the acquisition of the aspectual Form *–te i-(ru)* by learners of Japanese as a foreign language. *Language Learning* 54. 311–394.
Jacobsen, Wesley. 1992. *The transitive structure of events in Japanese*. Tokyo: Kurosio.
Kindaichi, Haruhiko. 1950. Kokugo Dooshi no ichibunrui [Classification of Japanese verbs]. *Gengo Kenkyuu* 15. 48–63.
Klein, Wolfgang. 2009. How time is encoded. In Wolfgang Klein and Ping Li (eds.), *The expression of cognitive categories*, 39–81. Berlin: Mouton de Gruyter.
Koyama, Satoru. 1998. Nihongo gakushūsha ni yoru tensu asupekuto no shūtoku [The acquisition of tense-aspect by learners of Japanese], *9th National Meeting of the Japanese Association of Second Language Acquisition*. Nagoya: Nagoya University.
Koyama, Satoru. 2004. Nihongo no tensu/asupekuto no shūtoku ni okeru fuhensei to kobetsusei: Bogo no yakuwari to eikyō o chuushin ni. [Universality and individuality in the acquisition of Japanese tense/aspect: Roles and effects of L1]. In Satoru Koyama, Kanako Ootomo and Miwako Nohara (eds.), *Gengo to kyōiku: Nihongo o taishō to shite.* [Language and education: Japanese as the target], 415–436. Tokyo: Kuroshio Publishers.
Kurono, Atsuko. 1995. Shokyū nihongo gakushūsha ni okeru –teiru no shūtoku ni tsuite [On the acquisition of *–teiru* by elementary learners of Japanese]. *Nihongo Kyōiku* [Journal of Japanese Language Teaching] 87. 153–164.
Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. In Ivan Sag and Anna Szabolsci (eds.), *Lexical matters*, 29–53. Stanford: CSLI.
Li, Ping and Yasuhiro Shirai. 2000. *The acquisition of lexical and grammatical aspect*. Berlin: Mouton de Gruyter.
Li, Ping and Sandra A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkely, CA: University of California Press.
Long, Shengyan, Hajime Ono and Hiromu Sakai. 2010. Bunrikai no katei ni okeru jishōtaipu no ninchi-Nihongo no asupekuto jōhōshori o tegakari ni [Recognition of event type in sentence comprehension: A view from processing of aspectual information in Japanese]. *Cognitive studies* 17(2). 313–331.
Long, Shengyan, Yuko Nakaishi, Hajime Ono and Hiromu Sakai. 2012. Dai 2 gengo no bunsyori ni doosi no asupekuto zyoohoo ga oyobosu eikyoo: Chuugokugo o bogo to suru nihongo gakusyuusya o taisyoo to site [Effects of aspectual information of verbs on L2 sentence processing: A study on Chinese learners of Japanese]. *Studies in Language Sciences* 11. 198–221.
Mackey, Alison and Susan Gass. 2012. *Research methods in second language acquisition: A practical guide*. Malden, MA: Wiley-Blackwell.

Mazurkewich, Irene. 1984. The acquisition of the dative alternation by second language learners and linguistic theory. *Language Learning* 34. 91–109.

McClure, William. 1995. *Syntactic projections of the semantics of aspect*. Tokyo: Hituzi Syobo.

Nishi Yumiko. 2008. *Verb learning and the acquisition of aspect: Rethinking the universality of lexical aspect and the significance of L1 transfer*. Ithaca, NY: Cornell University dissertation.

Ogihara, Toshiyuki. 1998. The ambiguity of the *–te iru* form in Japanese. *Journal of East Asian Linguistics* 7. 87–120.

Ogihara, Toshiyuki. 1999. Tense and aspect. In Natsuko Tsujimura (ed.), *The Handbook of Japanese linguistics*, 326–348. Malden, MA: Blackwell Publishing.

Olsen, Mari Broman. 1997. *A semantic and pragmatic model of lexical and grammatical aspect.* New York/London: Garland Publishin.

Sawasaki, Koichi and Akiko Kashiwagi-Wood. 2015. Issues in L2 Japanese sentence processing: similarities/differences with L1 and individual differences in working memory. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Sheu, Shiah-pey. 1997. Chū-jōkyū taiwanjin nihongo gakushūsha ni yori teiru no shūtokuni kansuru ōdan kenkyū [A cross-sectional study of the acquisition of *–te iru* by intermediate and advanced Taiwanese learners of Japanese]. *Nihongo Kyooiku* [Journal of Japanese Language Teaching] 95. 37–48.

Sheu, Shiah-pey. 2000. Shizen hatsuwa ni okeru nihongo gakushūsha ni yoru teiru no shūtoku kenkyū: OPI dēta no bunseki kara [Acquisition study of *–te iru* in natural speech by Japanese learners: Based on an analysis of OPI data]. *Nihongo Kyōiku* [Journal of Japanese Language Teaching] 104. 20–29.

Shibata, Miki. 1999. The use of Japanese tense-aspect morphology in L2 discourse narratives. *Acquisition of Japanese as a Second Language* 2. 68–102.

Shiokawa, Eriko. 2007. Nihongogakushūsha ni yoru asupekuto keishiki “te iru” no shūtoku – Bunmatsu to rentaishūshokusetsu tono kankei o chūshin ni. [Acquisition of the aspectual form *–te iru* by L2 learners of Japanese: Relation between main clauses and noun-modifying clauses]. *Nihongo Kyōiku* [Journal of Japanese Language Teaching] 134. 100–109.

Shirai, Yasuhiro. 1991. *Primacy of aspect in language acquisition: Simplified input and prototype.* Los Angeles, CA: University of California at Los Angeles dissertation.

Shirai, Yasuhiro. 1993. Inherent aspect and the acquisition of tense/aspect morphology in Japanese. In Heizo Nakajima and Yukio Otsu (eds.), *Argument structure: Its syntax and acquisition*, 185–211. Tokyo: Kaitakusha.

Shirai, Yasuhiro. 1998. The emergence of tense-aspect morphology is Japanese: Universal predisposition? *First Language* 18. 281–309.

Shirai, Yasuhiro. 2000. The semantics of the Japanese imperfective *–te iru*: An integrative approach. *Journal of Pragmatics* 32. 327–361.

Shirai, Yasuhiro. 2002a. The aspect hypothesis in SLA and the acquisition of Japanese. *Acquisition of Japanese as a Second Language* 5. 42–61.

Shirai, Yasuhiro. 2002b. The prototype hypothesis of tense-aspect acquisition in second language. In Rafael Salaberry and Yasuhiro Shirai (eds), *The L2 acquisition of tense-aspect morphology*, 451–474. Amsterdam: John Benjamins.

Shirai, Yasuhiro. 2004. A multiple-factor account to form-meaning connection in the acquisition of tense-aspect morphology. In Bill VanPatten, Jessica Williams, Susanne Rott and Mark Overstreet (eds.), *Form-meaning connections in second language acquisitions*, 91–112. Mahwah, NJ: L. Erlbaum Associates.

Shirai, Yasuhiro. 2015. L2 acquisition of Japanese. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Shirai, Yasuhiro and Atsuko Kurono. 1998. The acquisition of tense-aspect marking in Japanese as a second language. *Language Learning* 48. 245–279.

Shirai, Yasuhiro and Yumiko Nishi. 2005. How what we mean impacts how we talk: The Japanese imperfective marker te-iru in conversation. In Jan Frodesen and Christine Holten (eds.), *The power of context in language teaching and learning*, 39–48. Boston, MA: Heinle and Heinle.

Smith, Carlota. 1991. *The parameter of aspect*. Dordrecht: Kluwer.

Smith, Carlota. 1997. *The parameter of aspect (*2nd. Ed). Dordrecht: Kluwer.

Smith, Carlota. 2008. Time with and without Tense. In Jacqueline Gueron and Jacqueline Lecarme (eds.), *Time and modality*, 227–249. New York: Springer.

Sugaya, Natsue and Yasuhiro Shirai. 2007. The acquisition of progressive and resultative meanings of the imperfective aspect marker by L2 learners of Japanese. *Studies in Second Language Acquisition* 29. 1–38.

Sugita, Mamori. 2009. *Japanese* –te iru *and* –te aru*: The aspectual implications of the stage-level and individual-level distinction.* New York: City University of New York Graduate Center dissertation.

Tenny, Carol. 1994. *Aspectual roles and the syntax-semantics interface*. Dordrecht: Kluwer.

Valin, Robert van, Jr. 2006. Some universals of verb semantics. In Ricardo Mairal and Juana Gil (eds.), *Linguistics universals.* Cambridge: Cambridge University Press.

Vendler, Zeno. 1967. Verbs and times. In Zeno Vendler (ed.), *Linguistics in philosophy*, 97–121. Ithica, NY: Cornell University Press.

Verkuyl, Henk. 1993. *A theory of aspectuality: The interaction between temporal and atemporal structure.* New York: Cambridge University Press.

Hiroko Hagiwara

# 10 Language acquisition and brain development: Cortical processing of a foreign language

## 1 Introduction

Over the past decades the field of Cognitive Neuroscience, or the research on the language-brain relationship, has made great progress. New imaging techniques have allowed us to identify the neuronal network supporting language functions. Techniques providing a high temporal resolution such as event-related brain potentials (ERPs) and a high spatial resolution such as functional Magnetic Resonance Imaging (fMRI) have described the time course and brain regions of neuronal activities related to particular language functions, or modules, such as phonology, syntax, and semantics as well as the possible interplay among these modules. The combination of the new techniques with theoretical linguistic and psycholinguistic theorizing has clarified various aspects of language processing in normal adult brains. One of the key issues in current cognitive neuroscience is how language acquisition and brain development co-occur in early childhood and when cortical plasticity for these language modules deteriorates.

The notion of critical period, or sensitive period, for language acquisition comes from the loss of flexibility for cerebral reorganization due to acquired aphasia after puberty (Lenneberg 1967). This notion is extended to second language acquisition (L2) and it has been controversial whether complete mastery of a language is impossible after puberty. It is possible that the different linguistic modules are developed in different rates and that the timing and duration of their critical periods differ. For example, Weber-Fox and Neville (1996, 2001) suggested that syntactic aspects of sentence processing are more severely affected by the age of immersion in the L2 than semantic aspects. Furthermore, some researchers claim that knowledge of L2 attained after childhood is qualitatively different from that of a native or first language (L1) attained during childhood (Bley-Vronam 1990), and that highly proficient L2 learners differ from native speakers only quantitatively, but not qualitatively (Epstein, Flynn and Martohardjono 1996). In this chapter, we will focus on the L2 acquisition by Japanese learners of English both in adulthood and childhood. We will show that a certain aspect of syntax, core computation in narrow syntax and morphology-syntax interface is free from the notion of the critical period, and that word learning in childhood is biologically constrained in the human brain. Furthermore, cortical processing of words in childhood will be explored with respect to phonological and semantic aspects. It is important to note that this chapter looks at

the status of L2 in the Japanese brain. The organization of this chapter is as follows: We will discuss the status of L2 morphosyntactic issues in adult Japanese learners in the next section, and then, the status of L2 words in Japanese elementary school children's brains in section 3. Finally, section 4 provides our concluding remarks.

## 2 Neural correlates of L2 acquisition of English in adult Japanese

It is well known that even advanced or early L2 learners use inflectional morphology variably under circumstances in which native speakers obligatorily use it (Hawkins and Chan 1997; Ionin and Wexler 2002; Prévost and White 2000; see also Nakayama and Yoshimura's chapter in this volume). There are two main perspectives in L2 research to account for L2 learners' morphological variability. The first perspective concerns full access to the Universal Grammar (UG) position, provided that UG is characterized as the initial state in the faculty of language and has been regarded as a mental organ (Chomsky 2007). In this perspective, L2 learners' morphological variability is UG constrained and is assumed not to be a competence problem but rather a *performance* problem, the core notion of which was stated as in the Missing Surface Inflection Hypothesis (MSIH) (Prévost and White 2000). The second perspective stands for partial access to the UG position, where L2 learners' morphological variability is assumed to be due to a deficiency in acquiring uninterpretable features that are not present in L1 at the *computational level*, typically referred to as the Representational Deficit Hypothesis (RDH) (Hawkins 2005). Uninterpretable features are void of semantic content but are crucial for syntactic representations (e.g. phi (φ) features of verbs ([number]/[person]/[gender])), and this contrasts with interpretable features (e.g. tense feature of verbs ([present]/[past])). It predicts that uninterpretable features that have not been selected during the critical period are not available.

Some hypotheses and related studies suggested that L2 learners' morphological variability is related to performance factors, such as communication pressure (Ionin and Wexler 2002; Prévost and White 2000). However, elimination of the performance factors during speech production and behavioral experiments is difficult. In this context, the use of the ERP technique is ideal to minimize articulatory performance errors. The physiological and psychological status of cognitive processing of linguistic stimuli are directly reflected in the brain responses of neuronal activities and its characteristics of high-temporal resolution let us conduct a fine-grained analysis on a millisecond basis. Furthermore, neurophysiological indexes such as N400 (negativity observed around 400 milliseconds) for semantic processing, LAN (left anterior negativity) for morphosyntactic processing and P600 (positivity observed around 600 milliseconds) for repair and reanalysis for syntactic processing, have been repeatedly reported in many languages, e.g. English (Kutas and Hillyard 1980; Neville, Nicol,

Barss, Forster and Garrett 1991; Kutas and Federmeier 2011), German (Friederici, Pfeifer and Hahne 1993; Friederici 2002), Dutch (Hagoort, Brown and Groothusen 1993) and Japanese (Nakagome et al. 2001; Hagiwara et al. 2000; Koso, Ojima and Hagiwara 2011) to mention just a few. Therefore, it is possible to investigate L2 learners' sensitivity to L2 morphosyntactic violations and the underlying neural mechanisms associated with their sensitivity (See Sakamoto, this volume, for more details on ERP studies on language processing).

It is well known that L2 processing is quantitatively and/or qualitatively different from L1 processing and is modulated by factors such as the age of L2 acquisition, L2 proficiency level, and the linguistic properties of L1 known as L1 transfer. Using ERPs in L2 research, we can detect quantitative (the relative degree of latency or the amplitude of the component) and qualitative (presence or absence of the component or distinct polarity or topography) differences in the time course and of the degree of neuronal activity during language processing among the populations of different language backgrounds.

## 2.1 ERP studies of semantic and morphosyntactic processing in adult Japanese learners of English: Effects of proficiency

Only a few studies have used ERP experiments to investigate the neural mechanisms underlying English morphosyntactic processing in Japanese learners of English (JLEs). In the ERP study by Ojima, Nakata, and Kakigi (2005), English stimuli comprising the conditions of semantic expectations as in (1) and those of subject-verb agreement such as in (2) were visually presented to late JLEs, who acquired their L2 after the age of 12, with high or low English proficiency and to native English speakers (ENG).

(1) *The house has ten* rooms/*cities *in total.*

(2) *Turtles* move/*moves *slowly.*

The results indicate that semantically incongruent sentences (1) elicited N400 both in high and low groups of JLEs and ENG, compared to the congruent sentences, suggesting that L2 semantic processing is acquired even in the low proficiency group. On the other hand, the sentences with the violation of subject-verb agreement as in (2), compared to the well-formed sentences, elicited both LAN and P600 in ENG, while only LAN was observed in the high group of JLEs. Neither LAN nor P600 was observed in the low group. In other words, L2 morphosyntactic processing in the late JLEs was close to the native-like neural responses with high L2 proficiency, as shown by the appearance of LAN elicited in the high proficiency group. Based on this fact, Ojima et al. have argued against a critical period hypothesis,

which claims a fundamental difference between post-childhood L2 learning and childhood L1 learning. In addition to the effect of the L2 proficiency level, they have argued for the effect of L1 transfer because linguistic features required for subject-verb agreement in English are not present in Japanese.

Wakabayashi et al. (2007) also conducted an ERP experiment with late intermediate JLEs who acquired their L2 after the age of 12. The results showed that P600 was observed in the sentences with the violation of subject-verb agreement in person such as in (3), but no P600 was evoked in the sentences of the subject-verb agreement violation in number such as in (4).

(3) *I* answer/*answers *your letter.*

(4) *The teachers* answer/*answers *our questions.*

Wakabayashi et al. explained the JLEs' insensitivity to subject-verb disagreement in number such that the problem was not only due to mapping from syntax to morphology but also to the [number] feature. Since the [number] feature is not present in Japanese in numeration, and was also an optional feature, which is specified by operations in numeration, JLSs have to newly learn this feature (See also Nakayama and Yoshimura, this volume, Section 2.1 on behavioral studies of L2 acquisition).

The results of these studies indicate that the levels of proficiency and the linguistic properties of L1 modulate the characteristics of the ERP component, suggesting the importance of these factors in L2 acquisition. However, the neural mechanisms underlying the processing of L2 in early JLEs who started learning English before the age of 12, i.e., before they entered junior high school, have not been clarified. It is still unclear whether the age of acquisition plays a crucial role in L2 acquisition in foreign language context.

Tatsuta and Hagiwara (2012) conducted an ERP experiment using a high-density EEG system (128 channels), in which JLEs were divided into groups on the basis of their age when they began to learn English (*Early* or *Late*) as well as their English proficiency level (*High* or *Low*). Stimulus sentences were also designed to examine the effect of L1 transfer on English processing. The materials consisted of English stimuli for Present (subject-verb agreement in number) and Past (past tense inflection) conditions. One type of Present condition had a quantifier or a numeral in front of the subject determiner phrase for characterizing the plurality of the subject in a sentence, as shown in (5), and the other type did not have either of them, as shown in (6). In the Past condition a past tense adverb phrase was placed in front of a main clause, and verb was inflected for the past tense, as shown in (7) and (8), which were modified versions of (5) and (6), respectively.

(5) *Many boys* like/*likes *movies with action.*

(6) *Every evening, the little sisters* help/*helps *their mother.*

(7) *In those days, many boys* liked/*like *movies with action.*

(8) *Last night, the little sisters* helped/*help *their mother.*

The learners were divided into a group of JLEs who started learning English before the age of 12 (*Early* group) and a group of JLEs who started learning after the age of 12 (*Late* group). Furthermore, each group was subdivided into a group with high English proficiency (*High* group) or a group with low English proficiency (*Low* group). Accordingly, there were four JLEs groups: (a) the *Early-High* (EH) group, (b) the *Early-Low* (EL) group, (c) the *Late-High* (LH) group and (d) the *Late-Low* (LL) group (Table 1). In addition, native English speakers (ENG) (n = 17; Mean age = 25.71; age range: 19-30 years) participated as a control group.

**Table 1:** Characteristics for each JLEs group in Tatsuta and Hagiwara (2012)

| Group | *No. of participants* (women) | Age *M* (*SD*) | *Oxford Placement Test*[a] *M* (*SD*) |
|---|---|---|---|
| EH (*Early-High*) | 23 (14) | 22.63 (3.32) | 45.62 (4.63) |
| EL (*Early-Low*) | 20 (12) | 22.94 (4.73) | 25.33 (4.54) |
| LH (*Late-High*) | 23 (11) | 24.45 (3.24) | 46.74 (3.32) |
| LL (*Late-Low*) | 21 (9) | 22.31 (4.55) | 26.92 (5.23) |

[a] The *Quick Placement Test* is a written multiple-choice test that has 60 questions on English morphosyntax and the scores range from 0 to 60 (University of Cambridge Local Examinations Syndicate [UCLES], 2001). The scores in the *High* groups (EH and LH) ranged from 40 to 54 and that in the *Low* groups (EL and LL) ranged from 18 to 39.

**Table 2:** The summary of the ERP results

| | Present condition Subject-verb agreement in number) | | Past condition past tense inflection) | |
|---|---|---|---|---|
| Group | *LAN* | *P600* | *LAN* | *P600* |
| ENG | ✓ | ✓ | ✓ | ✓ |
| EH (*Early-High*) | – | ✓ | early negativity | ✓ |
| EL (*Early-Low*) | – | – | early negativity | – |
| LH (*Late-High*) | ✓ | ✓ | sustained negativity | – |
| LL (*Late-Low*) | – | – | sustained negativity | – |

NB: ✓: the component was observed; – : no ERP component was observed.

The summary of the ERP results is shown in Table 2. The ERP results for the Present condition in the ENG group showed a typical ERP pattern for morphosyntactic processing, namely a biphasic pattern with LAN followed by P600. Among

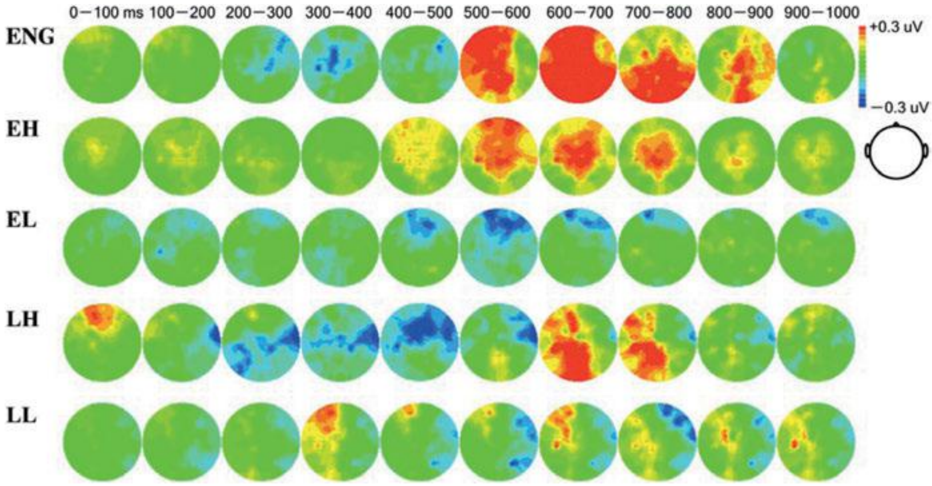



**Figure 1:** Scalp topographies for the Present condition (subject-verb agreement) in each group: The red color indicates positivity and the blue color shows negativity.

those in the JLE groups, the *Early-High* group showed P600, the *Late-High* group exhibited negativity with a broad distribution from 300 to 500 ms after the stimulus (LAN) followed by P600, and no ERP component was observed in the two *Low* groups. The scalp topographies of the Present condition for each group are shown in Figure 1.

These results exhibit the following four characteristics: First, although the negativity was quantitatively different in the onset latency (ENG, 400–450 ms; LH, 200–300 ms) and in the distribution from the ENG group, the *Late-High* group exhibited similar ERP components to the ENG group, i.e., a LAN and a P600. These results, together with the results of the *Early-Low* group of no responses, suggest that the age factor alone does not play a crucial role in the acquisition of English by the JLEs. Second, the *Early-High* group exhibited only P600. Despite the different ERP patterns from the ENG, the appearance of P600 without LAN in the *Early-High* group replicated the results found in high L2 learners in previous studies (Hahne 2001; Rossi, Gugler, Friederici and Hahne 2006; Tokowicz and MacWhinney 2005). The authors interpreted these results such that the *Early-High* group was able to perform the native-like mechanisms of the P600-indexed late controlled morphosyntactic processing, and that morphosyntactic marking, which was supposed to be represented by the appearance of LAN, for subject-verb agreement was less crucial for the assignment of grammaticality in a given sentence in the *Early-High* group than in the ENG and *Late-High* groups. Third, no ERP component in the *Low* groups, which also replicated the results obtained in lower L2 learners in previous studies (Hahne 2001; Ojima et al. 2005), suggests that the *Low* groups did not process the

operation of subject-verb agreement in English. The lack of LAN and P600 in the two *Low* groups could have been due to the effect of L1 transfer of the morphological representation systems, as there is no agreement in number in Japanese, in contrast to English. Finally, P600 in the two *High* groups and no ERP component in the two *Low* groups showed an effect of the English proficiency level on the P600-indexed controlled morphosyntactic processing, suggesting that JLEs could process the operation of subject-verb agreement once their English proficiency reached a higher level.

The ERP results in the Past condition, on the other hand, showed different patterns from those of the Present condition with respect to the JLE groups, while the ERP pattern of the ENG group remained the same as that in the Present condition. The two *Early* groups showed an early negativity from 300 to 500 ms in a mid anterior region and the two *Late* groups showed a sustained negativity from 300 to 800 ms around the anterior part of the scalp. These results suggest that JLEs were truly sensitive to past tense inflection in English, and that the sensitivity appeared to be qualitatively different from that in ENG group. Also, the different types of negativities observed differently in the two *Early* groups and the two *Late* groups suggest the effect of age on learning English. Although reasons for such a difference in latency and scalp distribution of two negativities remain unknown, the authors interpreted the sustained anterior negativity in the two *Late* groups as a reflection of working memory for syntactic processing, which is widely observed in the processing of filler-gap dependencies in the sentences with *Wh- and NP-*movement (Fiebach, Schlesewsky and Friederici 2002; Kluender and Kutas 1993; Hagiwara, Soshi, Ishihara and Imanaka 2007). In processing sentences in the Past condition, verbal working memory for sentence comprehension is required because it involves a dependency between the past tense adverbial phrase and a verb with the tense feature, both of which are not adjacent to each other unlike the subject and a verb in the sentences of the Present condition. Interestingly, some studies reported that verbal working memory capacity could be an indicator for predicting the achievement of L2 acquisition (Harrington and Sawyer 1992).

In short, Tatsuta and Hagiwara (2012) showed that the L2 proficiency level and L1 transfer of the morphological representation systems affected the neural mechanisms underlying L2 morphosyntactic processing. With respect to the accessibility hypotheses, this study does not support the Representational Deficit Hypothesis (Hawkins 2005), because the results indicate that the JLEs who have achieved higher English proficiency were able to process the operation of subject-verb agreement.[1] The native-like brain activities in the *Late-High* group suggest that L2 processing was constrained by UG, which then supports the full access to the UG position, and

**1** The result of this study is immune to the Missing Surface Inflection Hypothesis (Prevost and White, 2000), which predicts the separation of syntactic representations from their phonological exponents in L2 learners, because visual stimuli, not auditory, were employed in this study.

argue against a critical period hypothesis. This study also cautions against interpreting native-like performance as evidence that there are no qualitative differences in the processing between native speakers and L2 learners, and against interpreting the same behavioral performance as evidence that there are no differences in the neural mechanisms of the processing among groups of L2 learners.

## 2.2 fMRI study of structural dependence in adult Japanese learners of English

As we have seen above, previous studies of the acquisition of L2 syntax have mainly focused on inflectional morphology or morphosyntax such as subject-verb agreement in number, gender, person, or tense. These linguistic phenomena show parametric variation across languages. It is noteworthy that the agreement obeys a universal syntactic constraint (e.g., locality), but the existence of overt morphology is not universal (Hale 1996). In this context, it is necessary to examine the principle that is purely universal, i.e., core computational aspect of syntax, which is nonparametric abstract knowledge of UG and is biologically constrained (Chomsky 2007). The principle of structural dependence is one such candidate. Previous behavioral studies tested the principle in L1 acquisition of English-speaking children (Crain and Nakayama 1987), and the inability to acquire invented (not natural) language by a polyglot savant (Smith and Tsimpli 1996), reaching the conclusion that it cannot be inferred from the input, and therefore, forms part of the human language recipe. However, no studies have examined the neural correlates of the principle in L2 acquisition.[2]

Yusa et al. (2011) investigated whether L2 learners' knowledge would go beyond the input or stimuli that they had received during instruction, by examining the acquisition of a syntactic rule called negative inversion (NI).[3] The rule of NI obeys the rule of structure dependence in that in a negation sentence, negative adverbs (*never, seldom, rarely,* etc.), when placed at the beginning of a sentence, obligatorily trigger inversion and must be followed by auxiliaries (*be*-verbs, *can, must, may,* etc.): *I will never eat sushi* → *Never will I eat sushi*. In simple sentences, the sentence (11) is formed from sentence (9) either by the rule of NI or by moving the first or left-most auxiliary after the fronted negative adverb (the structure-independent rule). In complex sentences, on the other hand, sentence (15) can be formed successfully from sentence (13) only by the rule of NI. The structure-independent rule wrongly produces sentence (16) from (13).

---

**2** See Nakayama and Yoshimura in this volume for behavioral L2 studies on the core computational aspect of syntax.

**3** It should be noted that NI is acquired late in L1 (Sobin, 2003) and the frequency of occurrence of the structure in native English input is relatively low. In addition, Japanese language does not have NI, which automatically rejects the possibility of L1 transfer in Japanese learners' acquisition of English.

(9) *Those students are never late for class.*

(10) **Those students are late for never class.*

(11) *Never are those students late for class.*

(12) **Never those students are late for class.*

(13) *Those students who are very smart are never silent in class.*

(14) **Those never students who are very smart are silent in class.*

(15) *Never are those students who are very smart silent in class.*

(16) **Never are those students who very smart are silent in class.*

Yusa et al. conducted an fMRI study with 40 adult Japanese learners of English, who were divided into two groups: 20 participants received instruction for about a month on NI only simplex sentences (9)–(12) and the other 20 participants received no instruction, in the context of other things being equal.[4] As for complex sentences such as (13)–(16), no instruction was given to the instruction group. The functional MRI experiment was conducted twice for each group, before the instruction (Test 1) and after the instruction (Test 2). During the MRI, participants were asked to judge whether the sentences presented visually were correct or not.

The results of the grammaticality judgment showed that, in the instruction group, the error rate decreased significantly after the instruction, compared to prior to the instruction, not only in the simplex sentences but also in the complex sentences for which they did not receive any instruction, suggesting that they acquired knowledge of NI after one month of instruction. As expected, no improvement in accuracy was obtained for the non-instruction group. Concerning the fMRI data, significant activation was observed only for the instruction group on the inversion conditions not only for simplex sentences but also for complex sentences after instruction. Figure 2 illustrates the cortical activation pattern of the left inferior frontal gyrus (IFG), a Broca's area where the most significant change was observed between comprehending complex sentences in Test 1 (before instruction) and in Test 2 (after instruction) of the instruction group (INS) in the grammatical inversion condition (A2) as well as in the ungrammatical inversion condition (B2).[5] In the non-instruction group (NON-INS), on the other hand, there was no significant change

---

**4** Two groups were considered to be qualitatively comparable in English knowledge at the time of the first fMIR measurement. There were no significant differences between the two groups in the mean scores on the TOEIC and error rates for the first fMRI scan. The mean age of first exposure to English in the instruction group was 12.4 ± 0.4 years and 12.5 ± 0.3 years in the non-instruction group. Participants in the instruction group met twice a week for one month (8 classes in total), with one training session lasting an hour in addition to their regular classes at the university. They were required to hand in assignments based on the training sessions and feedback was given to them.

**5** As for simplex sentences in the instruction group, no significant activation change (A1) or decrease of the activation (B1) had occurred after the instruction, suggesting the consolidation stage of the rule of NI as a result of explicit instruction (Cf. Indefrey 2006).

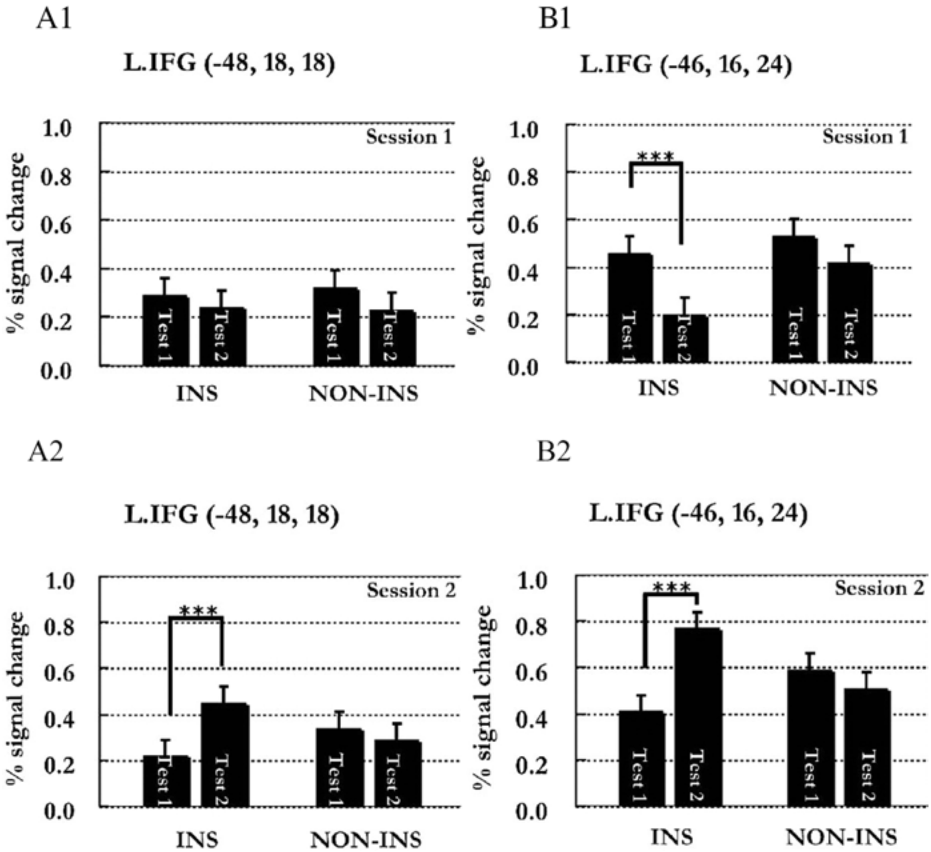



**Figure 2:** Brain activation for the pars triangularis of the left inferior frontal gyrus (LIFG), i.e., a part of the Broca's area. A1: sentence type (11), B1: sentence type (12), A2: sentence type (15), B2; sentence type (16), INS: instructed group, NON-INS: uninstructed group (Modified from Figure 4 from Yusa et al. (2011).)

between Test 1 and Test 2 in any of the sessions run. The left IFG is the area that the previous studies had identified as being responsible for the acquisition of a new rule (Musso et al. 2003; Tettamanti et al. 2002). Based on the results of neuroimaging data together with behavioral performance, Yusa et al. interpreted that the knowledge of the new rule NI which was amplified by the instruction of simplex sentences was conjectured to have projected into a rich knowledge of complex NI sentences that L2 learners have not been taught. They also claimed that the principle of structure dependence, one of the core principles of UG, still functions in L2 acquisition and makes it possible for L2 learners to know more than what is taught, which strongly argues against the critical period hypothesis.

In this section, we have seen some of the latest neurolinguistic studies of L2 acquisition of English by adult Japanese. We have found that the level of proficiency

is the most crucial factor in the investigation of L2 acquisition. Furthermore, contrary to the previous notion of the critical period, these studies have clearly demonstrated that the age of L2 acquisition does not affect the processing of subject-verb agreement in number nor a core syntactic principle of structure dependence. In other words, both the parametric rule (morphology-syntax interface) and nonparametric abstract knowledge of UG (narrow syntax) still functions after the critical period, which suggests plasticity for the neural circuits for language in adult L2 learners.

# 3 Neural correlates of foreign language learning in childhood: A cohort study

While neural mechanisms of phonology, morphosyntax, syntax and semantics in adults have been understood relatively well during the past few decades, those for normally developing children have not been fully investigated. One of the main reasons for this is the limitation in the use of neuroimaging techniques with high spatial resolution such as positron emission tomography (PET), functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG) on children. As PET uses injections of radioactive substance and fMRI and MEG use strong magnetic fields, their safety on the developing brain has not been proven to be safe. These instruments are also physically restrictive and unable to tolerate motion artifacts, and therefore, are not suitable for young children.

Near infrared spectroscopy (NIRS) is a relatively new technique that overcome these limitations. It can detect cerebral blood flow changes that are induced by neural activities, as signal changes of near infrared absorption through a concentration change in oxygenated- and deoxygenated hemoglobin. The major advantages include that it is fully non-invasive, unrestrictive and quiet, compared to PET and fMRI. Its components and setup are compact and measurement probes can be attached quickly and easily. Furthermore, since it can tolerate articulation-induced motion artifacts, we can measure not only perception or comprehension but also speech production. Recently, functional NIRS has been demonstrated to be an effective tool for monitoring local hemodynamic changes in the brain, especially for infants and even neonates (Peña et al. 2003; Homae et al. 2006). Therefore, this is quite suitable in developmental studies with children, especially for large-scale studies together with ERPs that have been successfully used to study L1 acquisition during childhood (Hahne, Eckstein and Friederici 2004; Holcomb, Coffey and Neville 1992). In the following, both ERPs and fNIRS studies of children's acquisition of word processing, although they are conducted separately, will be discussed.

## 3.1 ERP studies of semantic comprehension of spoken words in Japanese elementary school children

Most of the previous studies on child language acquisition were devoted to cross-sectional investigation, and few, if any, were conducted using a longitudinal paradigm. No studies have combined neuroimaging tools with behavioral assessments in cohort study. Cohort study is a large-scale longitudinal study that enables us to see the developmental change of the function in question within an individual or group over time. This is especially important in the field of second language acquisition since other factors such as age of first exposure to L2, L2-learning tasks and environments are held constant, thereby elucidating the causes that contribute to this change. The followings are some of the results of the large-scale cohort study on Japanese children's foreign language (FL) learning which aimed to see the neural mechanisms of FL acquisition during the ages of 6 to 11. See also Takahashi et al. (2011) and Hidaka et al. (2012) for similar challenging attempts on preschoolers.

Using the ERP technique, Ojima, Nakamura, Matsuba-Kurita, Hoshino and Hagiwara (2011) investigated children's cortical processing of FL words to provide direct and comprehensive neuroimaging evidence on child FL learning. During the experiment, Japanese children passively listened to words that were either congruous or incongruous in meaning with the preceding picture context. Previous L1 acquisition studies using this paradigm indicated the four developmental stages: no N400 negativity (stage 1), broad negativity (stage 2), typical N400 (stage 3) and a posterior N400 with a late positive component (LPC) (stage 4) (Friedrich and Friederici 2004, 2005; Silva-Pereyra, Riera-Gaxiola and Kuhl 2005; Hahne et al. 2004; Juottonen, Revonsuo and Lang 1996). The hypothesis Ojima, Nakamura et al. tested was that child FL learning closely follows stages in L1 acquisition (Dulay, Burt and Krashen 1982).

Table 3 shows the details of the 201 Japanese children who were selected out of the 322 children who participated in the cohort study for all 3 years, and who moved to the ERP analyses. They were born to a native Japanese-speaking mother and had lived in Japan until the end of the study.[6] On the basis of the English proficiency test, they were divided into the four groups: Low, Medium, High, and little progress, which served as a control group. Stimuli consisted of 80 basic-level English words appropriate for Japanese children and 80 Japanese words with corresponding meanings. Words difficult to understand due to cultural differences were not used.

6 The participants of the cohort study had different levels of English proficiency as they had different levels of exposure to L2. Some public schools provided 45-min English lessons (11–35 school h/year), while others did not. The children who went to public schools that did not provide English lessons had been exposed to English through commercial language schools and/or home study where parents/caretakers provided their children with exposure to English using videos, CDs, and other learning materials. Some children who went to English immersion schools were also included.

Examples of the stimuli included *akatyan* 'baby', *kaban* 'bag', *kuma* 'bear', *tori* 'bird', *hon* 'book', *hako* 'box', *neko* 'cat', *tukue* 'desk', *isya* 'doctor', *kao* 'face', *mon* 'gate' *boosi* 'hat', *kagi* 'key', *happa* 'leaf', *tizu* 'map', *sinbun* 'newspaper', *momo* 'peach', *yubiwa* 'ring', *hituzi* 'sheep, *densya* 'train', and *mado* 'window' [mean word length: 526 msec in English and 528 msec in Japanese; mean log10 of print frequency per million: 1.661 in English and 1.592 counts in Japanese].[7]

**Table 3:** Details of 201 children who were selected for the 3-year ERP analyses

| Group | No. of children | Mean age | *English score*[a] | | *AOFE*[b] | *HOE*[c] | |
|---|---|---|---|---|---|---|---|
| | | | Year 1 | Year 3 | | Year 1 | Year 3 |
| Low | 53 | 7.70 | 44.99 | 67.78 | 5.559 | 37.5 | 97.5 |
| Medium | 55 | 7.98 | 59.52 | 80.43 | 3.819 | 248.2 | 380.7 |
| High | 53 | 8.26 | 92.95 | 99.06 | 2.594 | 1740 | 2672 |
| Control[d] | 40 | 7.80 | 57.73 | 55.53 | 5.336 | 39.75 | 68.25 |

NB: [a] Mean score of the English test specifically designed for the cohort study; [b] Mean age of first exposure; [c] median hours of exposure; [d] a group of participants who had little progress.

The results of the longitudinal changes in each group are summarized in Figure 3. In Year 1, while congruous and incongruous FL words did not differ in the Low proficiency group (A), a broad negativity was evoked in the Medium proficiency group (B), and an N400 in the High proficiency group (C). In Year 3, a broad negativity appeared in the Low group (E), an N400 was elicited in the Medium group (F) and an LPC was found in High group in addition to an N400 (G). Put another way, each of the ERP responses of Low, Medium, and High groups was advanced one stage forward, i.e., stage 1 to stage 2 in Low, 2 to 3 in Medium, and 3 to 4 in High, respectively, over the period of two years. Furthermore, interestingly enough, ERP responses of FL words in the High group (G) are similar to those of the natives (D) in that both exhibit the biphasic pattern of the LPC preceded by the N400. As predicted, the ERP responses to the FL words are compatible with ERPs at Stages 1 to 4 in L1 acquisition, namely, these four stages appeared both in L1 acquisition and FL learning in the exact same order. In other words, these results indicate that both pattern and outcome of the ERP correlates in child FL learning resemble those for L1 acquisition. Considering that there exist large environmental and learning task differences between L1 acquisition and FL learning, Ojima, Nakamura et al. interpreted these data as reflecting learner-internal factors, namely, the biological nature of the brain itself determines the normal course of child FL learning.

One of the main issues in L2/FL acquisition is to clarify what contributes most in its mastery, e.g., the age of acquisition, amount of exposure, L1 transfer, learning

**7** The print frequencies of the stimulus words were based on studies using the most standard corpuses of English (Kučera and Francis 1967) and Japanese (Amano and Kondo 2000).

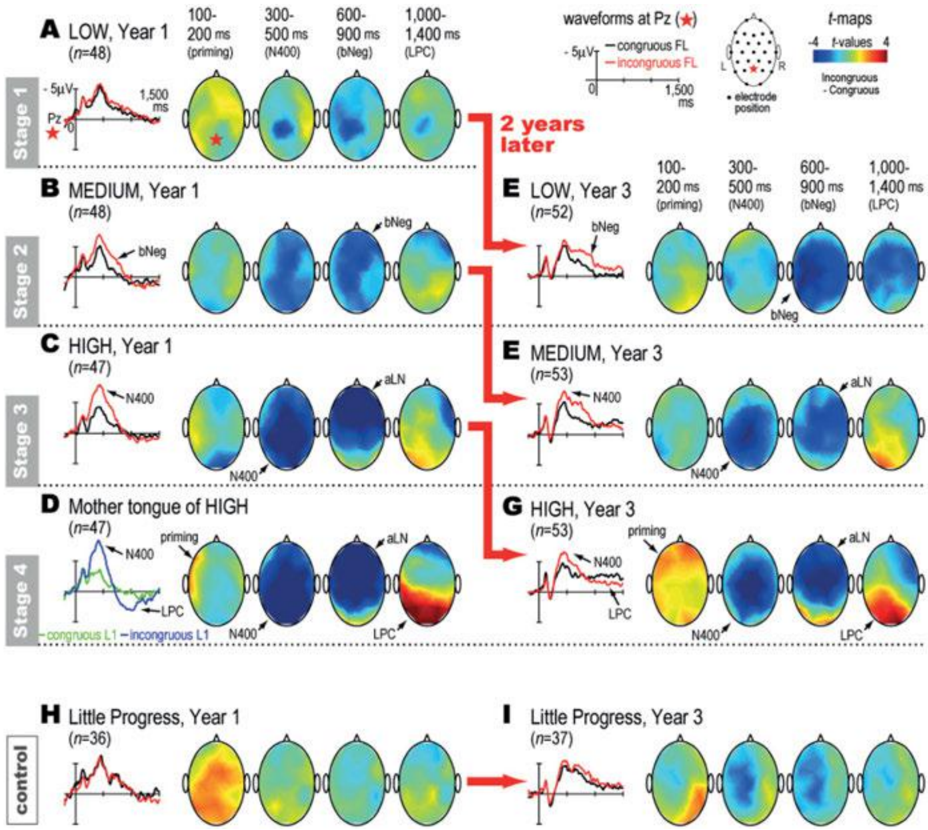



**Figure 3:** Longitudinal changes in ERP responses to English words by Japanese children (Reprinted from Ojima, Nakamura, Matsuba-Kurita, Hoshino and Hagiwara (2011) Neural correlates of foreign-language learning in childhood: A 3-year longitudinal ERP study. Journal of Cognitive Neuroscience 23. 183–199. Copyright © (2011) MIT Press. Reprinted with permission.)

environment, and learning strategy. Ojima, Matsuba-Kurita et al. (2011) focused on the age of first exposure (AOFE) and total hours of exposure (HOE). Based on a total of 815 ERP datasets obtained longitudinally from 350 children who participated in the cohort study, children's English proficiency scores and N400 amplitude were analyzed in multiple regression analyses.[8] The results showed that the effect of AOFE on the English score was significant, when that of log 10 HOE has been removed, indicating that *later*, rather than earlier, AOFE leads to higher English proficiency, when log 10 HOE is controlled for (Fig. 4A).[9] On the other hand, the effect of

**8** Multiple regression analyses can simultaneously assess each independent variable after controlling for the others (Zar 1999).

**9** As the index of amount of exposure, the common logarithm (log 10) of HOE rather than HOE itself was used. This is because HOE was related logarithmically, rather than linearly, to the English test score and the N400 amplitude.

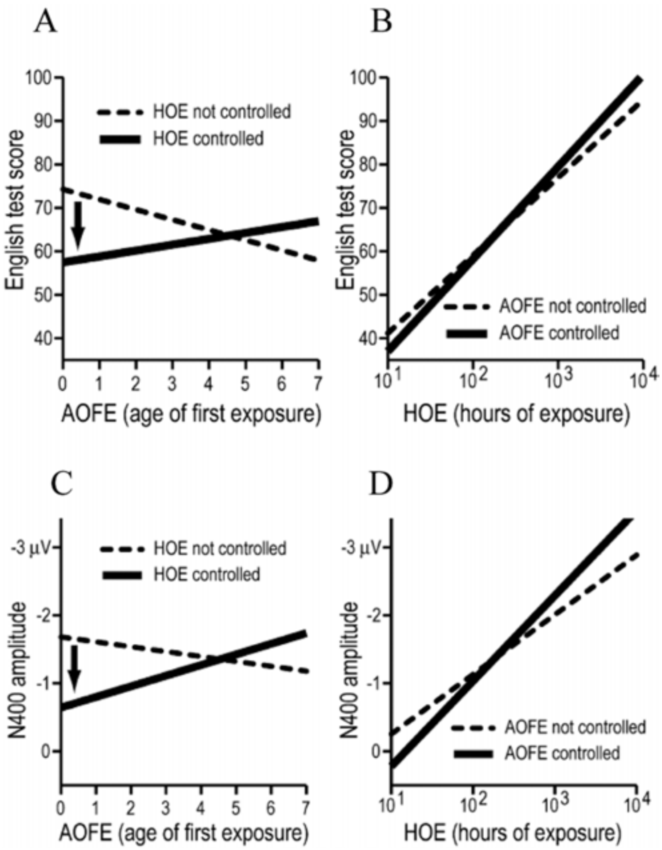



**Figure 4:** English test scores, mean N400 amplitudes and best-fitting regression lines (Modified from Figure 3 and Figure 5 from Ojima et al. (2011b).)

log 10 HOE on the English score was highly significant even after the effect of AOFE has been removed, showing that longer HOE leads to higher English proficiency, whether the AOFE is controlled or not (Fig. 4B).

The results of the ERP analyses also show that AOFE showed significant negative effects on the N400, when log 10HOE was controlled for (Fig. 4C). This means that children who had started English learning *later* showed larger N400 responses in English than did those who had started earlier and had had the same HOE. The effect of log 10 HOE on the N400 amplitude remained significant even after the effect of AOFE had been removed. Longer HOE leads to larger N400 amplitudes, whether AOFE had been controlled or not (Fig. 4D). On the basis of these results, the authors emphasized the importance of amount of exposure in FL learning, and cast doubt on the view that starting FL learning earlier always produces better results.

The results of these children's studies have some implications. First, the advantage of HOE over AOFE is somewhat unexpected because the effects of AOFE have

been reported repeatedly in previous behavioral studies (Yamada, Takatsuka, Kotake and Kurusu 1980: Johnson and Newport 1989; DeKeyser 2000), supporting the advantages of early starters in phonological skills and syntax learning. The discrepancy of the results in the previous behavioral studies and those in Ojima, Matsuba-Kurita et al. (2011) might be due to different phases or stages of learning. Previous behavioral studies on ongoing L2/FL learning have reported that late starters are faster learners than early starters even in syntax, suggesting the advantage of late starters in the speed of learning (Munoz 2006; Snow and Hoefnagel-Hohle 1978). This study dealt with children's *ongoing* FL learning as opposed to the adolescent and adults' *final outcome* of FL learning. Given that among the children who participated in the cohort study and learned English, the majority of them had been exposed to English at or before 7 years of age, so, a similar study should be conducted concerning AOFE beyond the age of 7. Needless to say, such a study must also include phonology and syntax learning.

Second, one of the effects critical for L2 other than AOFE and HOE would be the transfer from the mother tongue. We have already known that, in adults, L1 transfer of the morphological representation systems affect the neural mechanisms underlying L2 morphosyntactic processing (Wakabayashi et al. 1997; Tatsuta and Hagiwara 2012). As the ERP data in Ojima, Nakamura et al. (2011) and Ojima, Matsuba-Kurita et al. (2011) were obtained while children processed one word, future ERP research on L2 syntax must take into account possible effects of L1 transfer in the children and adolescents.

## 3.2 fNIRS investigation of phonological and semantic processing of words in child brain

Since we already know language skills continue to develop rapidly in children, we expect that brain structures and functions do so as well. A systematic observation of functional brain development in both L1 and L2 is crucial. While behavioral studies are abundant, there are only few studies dealing with normally developing children using neuroimaging techniques with high spatial resolution (Gaillard, Balsamo et al. 2003, Gaillard, Sachs et al. 2003; Sachs and Gaillard 2003; Szaflarski et al. 2006). Literature dealing with L2 acquisition is even more scarce, although studies dealing with older children have been conducted (Sakai et al. 2004; Tatsuno and Sakai 2005; Sakai et al. 2009).

Using fNIRS as a data acquisition tool and a basic word repetition task as a predictor of language learning ability, Sugiura et al. (2011) explored the different characteristics of language-related regions of interest (ROIs) and hemispheric laterality with respect to L1 and L2 processing of word frequency (high and low) in developing brains of school-age children. The participants are the same as those in the cohort study mentioned above, but the analysis is a cross-sectional examination

of the data obtained from the middle year of the cohort study. Among the 484 children (mean age: 8.93, SD: 0.89, age range: 6–10 years), the data of the 438 participants who were right-handed were analyzed in behavioral examination and fNIRS data of 392 participants were subjected to subsequent imaging analyses.

As for the experimental stimuli, a total of 120 single words were used: 30 words each for English high-frequency words, English low-frequency words, Japanese high-frequency words and Japanese low-frequency words.[10] The example of the stimuli for English high-frequency words included *brother, garden, picture, answer, become, carry, evening, pretty,* and *ready.* English low-frequency words were *fathom, nadir, schism, quorum, abash, cajole, devout, astute*, and *candid*. Japanese high-frequency words included *gaikoku* 'foreign country', *hookoku* 'report', *ningen* 'mankind', *ataeru* 'to give', *hazimeru* 'to begin', *kuraberu* 'to compare', *ookii* 'big', *saikoo* 'greatest' and *sukunai* 'few'. Japanese low-frequency words contained *tamamono* 'boon', *ibotei* 'half-brother', *adabana* 'abortive flower', *dokuduku* 'to abuse', *zunukeru* 'to exceed', *tunzaku* 'to burst through', *wabisii* 'dreary', *hagayui* 'impatient', and *azatoi* 'unscrupulous'.

Sugiura et al. tested the semantic knowledge of the word stimuli among the 4 repetition tasks and the results showed that mean semantic knowledge of the Japanese high frequency words (96%) was much higher than that of the Japanese low frequency words (12%), the English high frequency words (42%) and the English low-frequency words (8%). The comparison of the word repetition success rates between the 4 tasks showed significant differences in rates between all pairs (Jpn HF > Eng HF, Jpn LF > Eng LF, Jpn HF > Jpn LF, Eng HF > Eng LF, corrected $P < 0.001$).[11] High sensitivity of language familiarity in word repetition tasks indicates that language familiarity, not semantic knowledge, seems to be the crucial factor in the word repetition tasks.

The results of the brain activation pattern, at first, show that activation of L2 words were lower than that of L1 words in the superior/middle temporal gyrus, angular gyrus and supramarginal gyrus (Fig. 5 A, B, C). This fact suggests that L2 words were processed like nonword auditory stimuli in these cortical areas.

When one inspects language related regions closely, more specific characters of each region emerge with respect to phonological and semantic processing. In the superior/middle temporal gyri, during the repetition of L1 words, significantly greater activation was observed in the left hemisphere for high frequency words (96% semantic knowledge), whereas greater activation was observed in the right hemisphere for low frequency words (12% semantic knowledge) (Fig. 5A). These

**10** All Japanese words contained 4 moras and English words consisted of 2 syllables. The length of Japanese and English words was kept approximately equal. High-frequency words are defined as words that have >50 occurrences per million while the low-frequency words have <5 occurrences per million. All words used in the experiment were taken from Amano and Kondo (2000) for Japanese and Kučera and Francis (1967) for English.

**11** Whether the words were repeated correctly or not were evaluated phoneme-by-phoneme by a native Japanese who is a bilingual (English and Japanese) speaker.

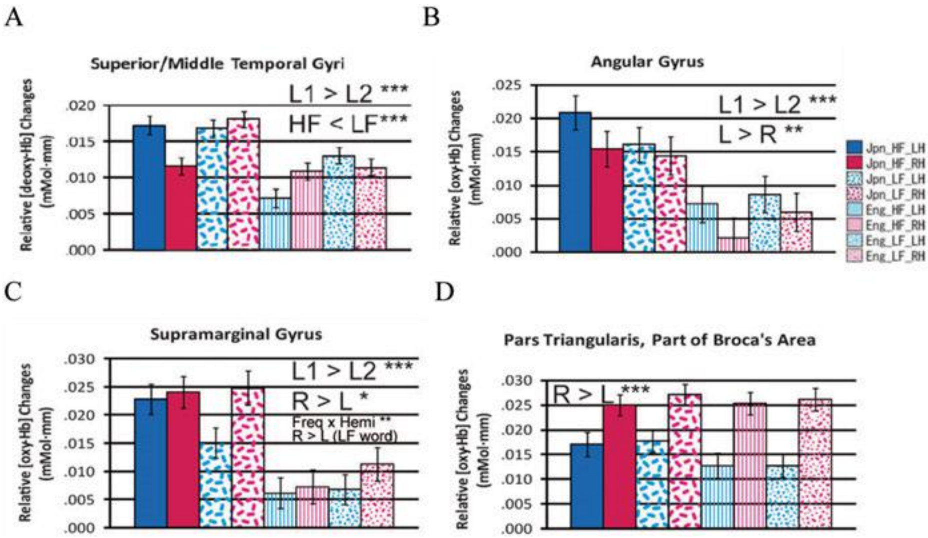



**Figure 5:** Average brain activation during word repetition tasks of 392 children. ROI analysis on the deoxy-hemoglobin [deoxy-HB] signals was shown in the superior/middle temporal gyri, including Wernicke's area (A); ROI analysis on the oxy-hemoglobin [oxy-HB] signals was shown in the angular gyrus (B), supramarginal gyrus (C) and pars triangularis, a part of Broca's area (D). L1: native language (Japanese); L2: second language (English). L: left hemisphere, R: right hemisphere. Jpn_HF_LH: Japanese high-frequency words in the left hemisphere; Jpn_HF_RH: Japanese high-frequency words in the right hemisphere; Jpn_LF_LH: Japanese low-frequency words in the left hemisphere; Jpn_LF_RH: Japanese low-frequency words in the right hemisphere; Eng_HF_LH: English high-frequency words in the left hemisphere; Eng_HF_RH: English high-frequency words in the right hemisphere; Eng_LF_LH: English low-frequency words in the left hemisphere; Eng_LF_RH: English low-frequency words in the right hemisphere. *: $P < 0.05$; **: $P < 0.01$; ***$P < 0.001$. (Modified from Figure 5 from Sugiura et al. (2011).)

results suggest that the left temporal region is engaged in semantic processing to some extent, whereas unknown words elicit more activation in the right hemisphere.

This hemispheric difference in phonological vs. semantic processing emerges more clearly in the angular gyrus and supramarginal gyrus. As Figures 5B and 5C show, low-frequency words of both L1 and L2 elicited more right-hemispheric activation in the supramarginal gyrus, whereas high-frequency words of both L1 and L2 elicited more left-hemispheric activation in the angular gyrus. These results suggest that the left angular gyrus is involved in semantic processing and the right supramarginal gyrus is involved in phonological processing.[12] Furthermore, the additional

12 The involvement of phonological processing in the supramarginal gyrus is a well- known fact from the lesion studies as well as MRI studies (Demonet, Price, Wise and Frackowiak 1994; Caplan, Gow and Makris 1995; Binder et al. 1996). This study newly found the bilateral right dominant activation in processing unfamiliar words in the supramarginal gyrus, suggesting phonological processing and storage in the right hemisphere.

involvement of phonological processing in the left angular gyrus could be observed from the comparison of the semantic knowledge of words, word repetition success rates and brain activation. While there was a significant difference in brain activation between L1 and L2 tasks, there was no significant difference between the children's semantic knowledge of L1 and L2 in low frequency words: L1 (12%) and L2 (8%). On the other hand, there was a significant difference in the word repetition success rates between L1 and L2 tasks both in high and low frequency words, which would reflect differences in phonological familiarity. Put it differently, in the left angular gyrus, processing familiar phonology in L1 induces higher brain activation than processing unfamiliar phonology in foreign language, independent of semantic knowledge. Moreover, these results suggest that a right-to-left shift in laterality occurs in the inferior parietal region as lexical knowledge increases, irrespective of language.

Significantly greater brain activation in the right hemisphere, compared to the left hemisphere, was also seen in the Broca's area (Fig. 5D). The involvement of phonological and prosodic processing in this region is supported by the fact that there were no differences in brain activation between high and low frequency word tasks nor a relationship between semantic knowledge and brain activation. Furthermore, previous studies reported that the role of the right Broca's area in prosodic processing has been demonstrated in pitch processing (Zatorre, Mondor and Evans 1999) and sentence melody processing (Meyer et al. 2002), and that the prosodic processing of the right hemisphere may facilitate the acquisition of lexical or syntactic knowledge in the early stages of language development (Homae et al. 2006). These processes would be equally valid for acquiring nonnative language. Sugiura et al. interpreted these results in the bilateral activation in Broca's area as presumed to be due to parallel processing, that is, left hemispheric segmental and right hemispheric suprasegmental information processing such as pitch, rhythm and intonation.

## 4 Concluding remarks

In summary, based on the ERP responses to the violation of subject-verb agreement in English by adult Japanese learners of English, we have found that the level of proficiency is the most crucial factor in investigating L2 acquisition. Processing sentences with negative inversion in English by adult JLEs demonstrated that the age of acquisition did not affect a core computational aspect of syntax, suggesting plasticity for neural circuits for language in adult L2 learners. A large-scale cohort study on elementary school children showed that L2 word processing in childhood is biologically constrained in the human brain, and that the hours of exposure, not the age of first exposure, accounted for L2 mastery. The functional NIRS brain imaging technique has revealed that the left angular gyrus is involved mainly in semantic

processing, and that the right supramarginal gyrus is involved in phonological processing. The greater involvement of the right Broca's area, compared to the left, in word processing suggests left hemispheric segmental and right hemispheric suprasegmental processing. Longitudinal cohort studies of developmental changes in brain function with respect to the acquisition of syntax in children are sought for future research.

## Acknowledgements

I would like to thank the editor of this volume Mineharu Nakayama for providing me with this opportunity to contribute to this volume. Much of the research cited in this chapter have been supported by a grant-in-aid for the promotion of "Brain Science and Education, Type II" from the Research Institute of Science and Technology for Society, Japan Science and Technology Agency (RISTEX/JST) to the author.

## References

Amano, Shigeaki and Tadahisa Kondo. 2000. *NTT database series: Nihongo-no Goitokusei* [Lexical properties of Japanese]. Tokyo: Sanseido.

Binder Jeffrey R., Julie A. Frost, Thomas A. Hammeke, Stephan M. Rao and Robert W. Cox. 1996. Function of the left planum temporale in auditory and linguistic processing. *Brain* 119. 1239–1247.

Bley-Vroman, Robert. 1990. The logical problem of foreign language learning. *Linguistic Analysis* 20. 3–49.

Caplan, David, David Gow and Nikos Makris. 1995. Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. Neurology 45. 293–298.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland and Hans-Martin Gärtner (eds), *Interfaces + Recursion = Language?*, 1–30. Berlin, Germany: Mouton de Gruyter.

Crain, Stephen and Mineharu Nakayama. 1987. Structure dependence in grammar formation. *Language* 63. 522–543.

DeKeyser, Robert M. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22. 499–533.

Demonet Jean-Francois, Cathy J. Price, Richard J. Wise and Richard S. J. Frackowiak. 1994. Differential activation of right and left posterior sylvian regions by semantic and phonological tasks: a positron-emission tomography study in normal human subjects. *Neuroscience Letters* 182. 25–28.

Dulay, H., M. Burt and Steven Krashen. 1982. *Language Two.* Oxford: Oxford University Press.

Epstein, Samuel David, Suzanne Flynn and Gita Martohardjono. 1996. Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences* 19. 677–758.

Fiebach, Christian J., Matthias Schlesewsky and Angela D. Friederici. 2002. Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German *WH* questions. *Journal of Memory and Language* 47. 250–272.

Friederici, Angela D. 2002. Towards a neural basis of auditory sentence processing. *Trends in Cognitive Science* 6. 78–84.

Friederici, Angela D., Erdmut Pfeiffer and Anja Hahne. 1993. Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research* 1. 183–192.

Friedrich, Manuela and Angela D. Friederici. 2004. N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience* 16. 1465–1477.

Friedrich, Manuela and Angela D. Friederici. 2005. Phonotactic knowledge and lexical–semantic processing in one-year-olds: Brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience* 17. 1785–1802.

Gaillard William Davis, Lyn M. Balsamo, Zuhaib Ibrahim, Bonnie C. Sachs and Ben Xu. 2003a. fMRI identifies regional specialization of neural networks for reading in young children. *Neurology* 60. 94–100.

Gaillard, William Davis, Bonnie C. Sachs, Joseph R. Whitnah, Zaaira Ahmad, Lyn M. McKinney, Kevin Hunter, Ben Xu and Cecile B. Grandin. 2003b. Developmental aspects of language processing: fMRI of verbal fluency in children and adults. *Human Brain Mapping* 18. 176–185.

Hagiwara, Hiroko, Heizo Nakajima, Kazuyuki Nakagome, Satoru Takazawa, Osamu Kanno, Kenji Itoh and Ichiro Koshida. 2000. ERP manifestations of processing syntactic dependencies in hierarchical structures of language: Time course and scalp distribution. In Kazuko Inoue (ed.) *Researching and verifying an advanced theory of human language*, 519–545. Chiba, Japan: Kanda University of International Studies.

Hagiwara, Hiroko, Takahiro Soshi, Masami Ishihara and Kuniyasu Imanaka. 2007. A topographical study on the event-related potential correlates of scrambled word order in Japanese complex sentences. *Journal of Cognitive Neuroscience* 19. 175–193.

Hagoort, Peter, Colin Brown and Joland Groothusen. 1993. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes* 8. 439–483.

Hahne, Anja. 2001. What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguist Research* 30. 251–266.

Hahne, Anja, Korinna Eckstein and Angela D. Friederici. 2004. Brain signatures of syntactic and semantic processes during children's language development. *Journal of Cognitive Neuroscience* 16. 1302–1318.

Hale, Kenneth. 1996. Can UG and L1 be distinguished in L2 acquisition? *Behavioral and Brain Sciences* 19. 728–730.

Harrington, Michael and Mark Sawyer. 1992. L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition* 14. 25–38.

Hawkins, Roger. 2005. Explaining full and partial success in the acquisition of second language grammatical properties. *Second Language* 4. 7–25.

Hawkins, Roger and Cecilia Yuet-hung Chan. 1997. The partial availability of universal grammar in second language acquisition: The 'failed functional features hypothesis.' *Second Language Research* 13. 187–226.

Hidaka, Souta, Hiroshi Shibata, Michiyo Kurihara, Akihiro Tanaka, Akitsugu Konno, Suguru Maruyama, Jiro Gyoba, Hiroko Hagiwara and Masatoshi Koizumi. 2012. Effect of second language exposure on brain activity for language processing among preschoolers. *Neuroscience Research* 73. 73–79.

Holcomb, Philip J., Sharon A. Coffey and Helen J. Neville. 1992. Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. *Developmental Neuropsychology* 8. 203–241.

Homae, Fumitaka, Hama Watanabe, Tamami Nakano, Kayo Asakawa and Gentaro Taga. 2006. The right hemisphere of sleeping infant perceives sentential prosody. *Neuroscience Reserch* 54. 276–280.

Ionin, Tanya and Kenneth Wexler 2002. Why is 'is' easier that '-s'?: Acquisition of tense/agreement morphology by child second language learning of English. *Second Language Research* 18. 95–136.

Indefrey, Peter. 2006. A meta-analysis of hemodynamic studies on first and second language processing: Which suggested differences can we trust and what do they mean? In Marianne Gullburg and Peter Indefrey (eds.) *The cognitive neuroscience of second language acquisition*, 279–304. Oxford, UK: Blackwell.

Johnson, Jacqueline S. and Elissa L. Newport. 1989. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21. 60–99.

Juottonen, Kirsi, Antti Revonsuo and Heikki Lang. 1996. Dissimilar age influences on two ERP waveforms (LPC and N400) reflecting semantic context effect. *Brain Research, Cognitive Brain Research* 4. 99–107.

Kluender, Robert and Martha Kutas. 1993. Bridging the gap: Evidence from ERPs on the processing unbounded dependencies. *Journal of Cognitive Neuroscience* 5. 196–214.

Koso, Ayumi, Shiro Ojima and Hiroko Hagiwara. 2011. An event-related potential investigation of lexical pitch-accent in auditory Japanese. *Brain Research* 1385. 217–228.

Kučera, Henry and W. Nelson Francis. 1967. *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Kutas, Marta and Steven A. Hillyard. 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207. 203–205.

Kutas, Marta and Kara D. Federmeier. 2011. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology* 62. 621–647.

Meyer, Martin, Kai Alter, Angela D. Friederici, Gabriele Lohmann and D. Yves von Cramon. 2002. Functional MRI reveals brain regions mediating slow prosodic manipulations of spoken sentences. *Human Brain Mapping* 17. 73–88.

Muñoz, Carmen (ed.) (2006). *Age and the rate of foreign language learning*, Clevedon, UK: Multilingual Matters.

Musso, Mariacristina, Andrea Moro, Volkmar Glauche, Michel Rijntjes, Jürgen Reichenbach, Christian Büchel and Cornelius Weiller. 2003. Broca's area and the language instinct. *Nature Neuroscience* 6. 774–781.

Nakagome, Kazuyuki, Satoru Takazawa, Osamu Kanno, Hiroko Hagiwara, Heizo Nakajima, Kenji Itoh and Ichiro Koshida. 2001. A topographical study of ERP correlates of semantic and syntactic violations in the Japanese language using the multichannel EEG system. *Psychophysiology* 38. 304–15.

Nakayama, Mineharu and Noriko Yoshimura. 2015. The modularity of grammar in L2 acquisition. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Neville, Helen, Janet L. Nicol, Andrew Barss, Kenneth I. Forster and Merrill F. Garrett. 1991. Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 3. 151–165.

Ojima, Shiro, Hiroki Nakata and Ryusuke Kakigi. 2005. An ERP study of second language learning after childhood: effects of proficiency. *Journal of Cognitive Neuroscience* 17. 1212–1228.

Ojima, Shiro, Naoko Nakamura, Hiroko Matsuba-Kurita, Takahiro Hoshino and Hiroko Hagiwara. 2011a. Neural correlates of foreign-language learning in childhood: a 3-year longitudinal ERP study. *Journal of Cognitive Neuroscience* 23. 183–199.

Ojima, Shiro, Hiroko Matsuba-Kurita, Naoko Nakamura, Takahiro Hoshino and Hiroko Hagiwara. 2011b. Age and amount of exposure to a foreign language during childhood: Behavioral and ERP data on the semantic comprehension of spoken English by Japanese children. *Neuroscience Research* 70. 197–205.

Peña Marcela, Atsushi Maki, Damir Kovačic, Ghislaine Dehaene-Lambertz, Hideaki Koizumi, Furio Bouquet, Jacques Mehler. 2003. Sounds and silence: an optical topography study of language recognition at birth. *Proceedings of National Academy of Sciences USA* 100. 11702–11705.

Prévost, Philippe and Lydia White. 2000. Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16. 103–133.

Rossi, Sonja, Manfred F. Gugler, Angela D. Friederici and Anja Hahne. 2006. The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience* 18. 2030–2048.

Sachs, Bonnie C. and William D. Gaillard. 2003. Organization of language networks in children: functional magnetic resonance imaging studies. *Current Neurology and Neuroscience Reports* 3. 157–162.

Sakai, Kuniyoshi L., Kunihiko Miura, Nobuko Narafu and Yukimasa Muraishi. 2004. Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cerebral Cortex* 14. 1233–1239.

Sakai, Kuniyoshi L., Arihito Nauchi, Yoshinori Tatsuno, Kazuyoshi Hirano, Yukimasa Muraishi, Masakazu Kimura, Mike Bostwick and Noriaki Yusa. 2009. Distinct roles of inferior frontal regions that explain individual differences in second language acquisition. *Human Brain Mapping* 30. 2440–2452.

Silva-Pereyra, Juan, Maritza Rivera-Gaxiola and Patricia K. Kuhl. 2005. An event-related brain potential study of sentence comprehension in preschoolers: Semantic and morphosyntactic processing. *Brain Research, Cognitive Brain Research* 23. 247–258.

Sobin, Nicholas. 2003. Negative inversion as nonmovement. *Syntax* 6. 183–212.

Smith, Neil and Ianthi-Maria Tsimpli. 1996. *The mind of a savant*, Oxford, UK: Blackwell.

Snow, Cathrine and Marian Hoefnagel-Hohle. 1978. The critical age for language acquisition: evidence from second language learning. *Child Development* 49. 1114–1128.

Sugiura, Lisa, Shiro Ojima, Hiroko Matsuba-Kurita, Ippeita Dan, Daisuke Tsuzuki, Takusige Katura, and Hiroko Hagiwara. 2011. Sound to language: Different cortical processing for first and second languages in elementary school children as revealed by a large-scale study using fNIRS. *Cerebral Cortex* 21. 2374–2393.

Szaflarski, Jerzy P., Vincent J. Schmithorst, Mekibib Altaye, Anna W. Byars, Jennifer Ret, Elena Plante and Scott K. Holland. 2006. A longitudinal functional magnetic resonance imaging study of language development in children 5 to 11 years old. *Annals of Neurology* 59. 796–807.

Takahashi, Junichi, Yuika Suzuki, Hiroshi Shibata, Yuichiro Fukumitsu, Jiro Gyoba, Hiroko Hagiwara and Masatoshi Koizumi. 2011. Effects of non-native language exposure on the semantic processing of native language in preschool children. *Neuroscience Researc*h 69. 246–251.

Tatsuno, Yoshinori and Kuniyoshi L. Sakai. 2005. Language-related activations in the left prefrontal regions are differentially modulated by age, proficiency, and task demands. *Journal of Neuroscience* 25. 1637–1644.

Tatsuta, Natsuko and Hiroko Hagiwara. 2012. English morphosyntactic processing of present and past tense forms in Japanese learners of English: An event-related brain potential study. Paper presented at the 38th annual conference of Japan Society of English Language Education, Aichi Gakuin University, Nagoya, 5 August.

Tettamanti, Marco, Hatem Alkadhi, Andrea Moro, Daniela Perani, Spyros Kollias and Dorothea Weniger. 2002. Neural correlates for the acquisition of natural language syntax. *Neuroimage* 17. 700–709.

Tokowicz, Natasha and Brian MacWhinney. 2005. Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition* 27. 173–204.

Wakabayashi, Shigenori, Kazuhiko Fukuda, Masanori Bannai and Shoichi Asaoka. 2007. Japanese speakers' sensitivity to third person singular *-s* in English: Arguments based on ERP data. *Second Language* 6. 19–46.

Weber-Fox, Christine and Helen J. Neville. 1996. Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience* 8. 231–256.

Weber-Fox, Christine and Helen J. Neville. 2001. Sensitive periods differentiate processing of open- and closed-class words: An ERP study of bilinguals. *Journal of Speech, Language and Hearing Research* 44. 1338–1353.

Yamada, Jun, Shigenobu Takatsuka, Nobuko Kotake, and Junko Kurusu. 1980. On the optimum age for teaching foreign vocabulary to children. *International Review of Applied Linguistics* 18(3). 245–247.

Yusa, Noriaki, Masatoshi Koizumi, Jungho Kim, Naoki Kimura, Shinya Uchida, Satoru Yokoyama, Naoki Miura, Ryuta Kawashima and Hiroko Hagiwara. 2011. Second-language instinct and instruction effects: Nature and nurture in second-language acquisition. *Journal of Cognitive Neuroscience* 23. 2716–2730.

Zar, Herrold H. 1999. *Biostatistical Analysis, 4th edition.* Upper Saddle River, NJ: Prentice Hall.

Zatorre, Robert J., Todd A. Mondor and Alan C. Evans. 1999. Auditory attention to space and frequency activates similar cerebral systems. *Neuroimage* 10. 544–554.

# II Japanese Language Processing

Yuki Hirose

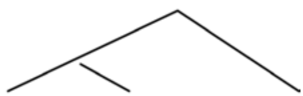
# 11 Resolution of branching ambiguity and the role of prosody

## 1 Introduction

In processing spoken sentences, studies have shown ample evidence that prosodic structure influences the choices that the listeners make in assigning syntactic structure to the incoming input (for review see Speer and Blodgett 2006; Cutler, Dahan and Donselaar 1997; Beckman 1996). Debates exist over whether certain prosodic cues are more informative depending on factors such as the task in which the speakers and the listeners are engaged in, or the referential ambiguity involved in the situation in syntactic processing. While some hold the view that the correspondence between syntax and the prosodic realization of the utterances is more or less constant, regardless of the task and the situation (Schafer, Speer and Warren 2005; Kraljic and Brennan 2005), others argue this relationship is dependent on whether or not the syntactic ambiguity can be referentially resolved (Snedeker and Trueswell 2003) or on the kind of task imposed on the speakers (Allbritton, McKoon and Ratcliff 1996). This is further complicated by the fact that there is no exact one-to-one mapping between the prosodic and the syntactic structure (Beckman 1996; Kubozono 1993; Selkirk 1984; Speer and Blodgett 2006; Shattuck-Hufnagel and Turk 1996). Recent findings suggest that speakers and listeners abide by different types of prosodic cues in encoding and decoding the syntactic structure (Kitagawa and Hirose 2012).

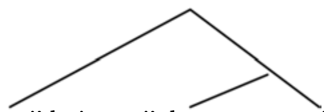
This chapter discusses the role of prosody in resolving the left- and right-branching ambiguity in Tokyo Japanese.

(1) a.



*mi'dori no* *i'nko no* *ma'huraa*
green parrot GEN scarf
'a scarf with a green parrot'

b.



*mi'dori no* *i'nko no* *ma'huraa*
green parrot GEN scarf
'a green scarf with a parrot'

Noun phrases such as *mi'dori no i'nko no ma'huraa* 'green parrot GEN scarf' (the accented mora in a word is marked by an apostrophe " ' " following it) are globally ambiguous as to whether the first element (color term + *no*) modifies the immediately following noun (N1) as in (1a) or the head of the entire noun phrase (N2) as in (1b) (In Japanese, color terms can either take an adjective form (e.g., *aoi*, 'blue') or the color name followed by the particle *no* (e.g., *midori no*, 'green')). From the

incremental processing perspectives, when N1 is first processed, N2 has not yet been encountered. Therefore, it is reasonable to assume that the modifier will be interpreted as being associated with N1. When the subsequent N2 is processed, the interpretation in which the modifier is actually attributed to N2 should only be achieved by reanalyzing the initial N1-modification interpretation. The same modifier-modificant ambiguity is present with other syntactic types of modifiers. In the case of processing relative clauses in Japanese, studies have found evidence for the local N1-interpretation (Kamide and Mitchell 1997; Miyamoto, Nakamura and Takahashi 2004). These results could also be accounted for by the demands of incremental processing (building structures immediately as the sentence unfolds, without waiting for subsequent information), that is, the N1 is encountered immediately following the relative clause in Japanese and at that point N1 is the only candidate for the head noun. Studies on processing Japanese relative clauses with this type of ambiguity actually produce somewhat mixed-results. For example, the reading times at the sentence-final region, and the off-line (after reading through the entire sentence) judgment data exhibit a tendency for N2-modification (=right branching structure) as the final decision. Apparently the final interpretation can override the initial analysis. See Kahraman and Sakai (this volume).

Back to the case of one-prosodic-word-long modifiers followed by N1 + N2, most evidence suggests that the phrase is preferentially interpreted with a left-branching structure (Ito, Arai and Hirose 2015). However, the extent to which the default preference can be modulated by various factors remains unclear. In this chapter we will consider the role of prosody in distinguishing the two alternative structures. In the following sections, in order to better control the factors other than prosody, we will focus on the cases in which the modifier (the first element) is at most one prosodic word, i.e., either an adjective or a genitive-marked NP, to start with.

## 2 Prosodic marking of the branching structure

Before going into discussion of how the two distinct branching structures are reflected in the prosodic structure, we should start by briefly describing word-level prosody in Tokyo Japanese. In Japanese, there is a contrast between accented and unaccented words at the lexical level. The position of the accent (i.e., the accented mora in the word) is also lexically determined (indicated by an apostrophe " ' " in this chapter). If a word is accented, the accented mora in it is associated with a sharp fall in fundamental frequency (F0), which is indicated by the accent tone H*+L in Tokyo Japanese.[1] An unaccented word starts with the initial pitch rising

1 For earlier works discussing how the accented (and unaccented) moras are specified for tones (H/L) in Japanese, see McCawley (1968), Haraguchi (1977) and Poser (1984).

(the L% followed by the phrasal H), but lacks the sharp F0 fall: instead, the F0 declines more gradually.

With respect to the perception of the word-level accent, Cutler and Otake (1996) and Otake and Cutler (1997) demonstrated that listeners were able to discriminate between different accent types exhibiting distinct tonal contours, for example, *ka'ge* (HL) 'shadow' vs. *kagi* (LH) 'key' by only being exposed to the initial mora in the gating task. Sugiyama (2012) examined native speakers' production and perception of pairs of bi-moraic nouns that were different in accent type, but are associated with the same surface tonal sequence. For example, *hana* 'nose' and *hana'* 'flower' contrast in accent type: the former is unaccented while the latter is accented on the final mora but they are realized with the same LH tone when pronounced in isolation. For the former, the LH contour consists of the boundary L tone followed by the phrasal (default) H tone; for the latter, it is the sequence of the boundary L tone followed by the accent H tone. Sugiyama reports that native speakers have difficulty distinguishing between the pairs when pronounced in isolation, where the L part of the accent tone H*L is not realized. As these studies together suggest, the detection of individual H and L tones associated with certain F0 ranges for the speaker appears to play a fundamental role in perception of Japanese lexical accent.

A pitch accent defines the intonation contour, which indicates the grouping of the constituents to form a higher level of phonological constituent. This in turn directly or indirectly corresponds to the syntactic structure. In the case of the above example in (1), all three elements are lexically accented. The prosodic phrase corresponding to each of the three elements, corresponding to each *bunsetsu* (a phonological phrase, in this case, an adjective, or noun phrases with case particles) in this case, is referred to by different terms depending on the theoretical framework: the *minor phrase* (mp) (Poser 1986; Selkirk 1986), or the *accentual phrase* (AP) (Beckman and Pierrehumbert 1986). The JToBI (Venditti 2005) and X-JToBI labeling scheme (Maekawa, Kikuchi, Igarashi and Venditti 2002; Igarashi, Kikuchi and Maekawa 2007) also use the latter term. For convenience, I will use "minor phrase" in the remainder of the chapter.

A minor phrase is associated with the initial pitch rising (of the L% boundary tone followed by the phrasal H tone) at the beginning of the word (*a.k.a.* initial lowering). The end of the phrase is again marked by the boundary L%. A minor phrase commonly consists of one or more *bunsetsu*, but can have one lexical accent at most. In Figure 1, I will cite the two contrasting F0 contours for two homophonic single minor phrases *uerumono* 'something to plant' / *ue'rumono* 'those who are starved', originally from Venditti (1994) (see also Venditti 2005; 2006; Venditti, Jun and Beckman 1996). In both cases, two words are grouped into a single minor phrase. They have the second unaccented word in common, but contrast in the accent type of the first word. In the left panel, the verb *ueru* 'to plant' is unaccented whereas in the right panel, the verb *ue'ru* 'to starve' is accented on the second mora.

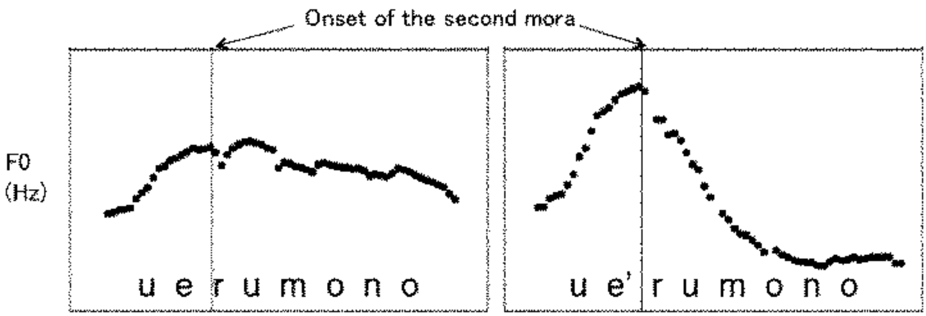



**Figure 1:** F0 contours for the accented phrase *uerumono* 'something to plant' (left) and *ue'rumono* 'those who are starved' (right), adapted from Venditti (2005)

See Venditti (2005, 2006) for discussions with more detailed examples of accent tones and phrasal tones in the J-ToBI schema.

In the sequence of minor phrases such as (1) above, we can expect downstep (also called catathesis), or the gradual declination in pitch range triggered by an accented element (Poser 1984; Pierrehumbert and Beckman 1988; Kubozono 1988). Downstep applies within a prosodic phrase dominating minor phrases. This level of prosodic phrase, defined as the domain of downstep, and which is to be dominated by *utterance* (the highest level in the prosodic hierarchy) is called *major phrase* (MP) (Poser 1984; Selkirk 1986), or the *intermediate phrase* (Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988). JToBI (Venditti 2005) maintains the latter term based on Beckman and Pierrehumbert (1986) and Pierrehumbert and Beckman (1988). The recent X-JToBI labeling scheme (Maekawa, Kikuchi, Igarashi and Venditti 2002; Igarashi, Kikuchi, and Maekawa 2007) instead uses intonation phrase (IP) as the highest level in the prosodic hierarchy (= utterance) which directly dominates accentual phrases.

In the case illustrated in Figure 2 (next page), downstep is applied over the three minor phrases, indicating the absence of a major phrase boundary intervening these elements. All three elements are within the same major phrase as in (2).

(2) $\{_{\text{MajP/ip/IP}}$ *mi'dori no i'nko no ma'furaa* }

However, in the right-branching structure, prosody demarcates the non-default structure by the raising of the pitch of the second element, apparently counteracting downstep, as shown in Figure 3 (next page). One interpretation of the phenomenon is that the downstep is reset and it starts a new major phrase (or an intermediate phrase, Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988). Selkirk and Tateishi (1991) (see also Selkirk 2000) explain this as a demand by the alignment theory between the prosodic phrasing and the syntactic phrasing: the beginning of a new syntactic phrase (a maximal projection level, in particular) and the beginning

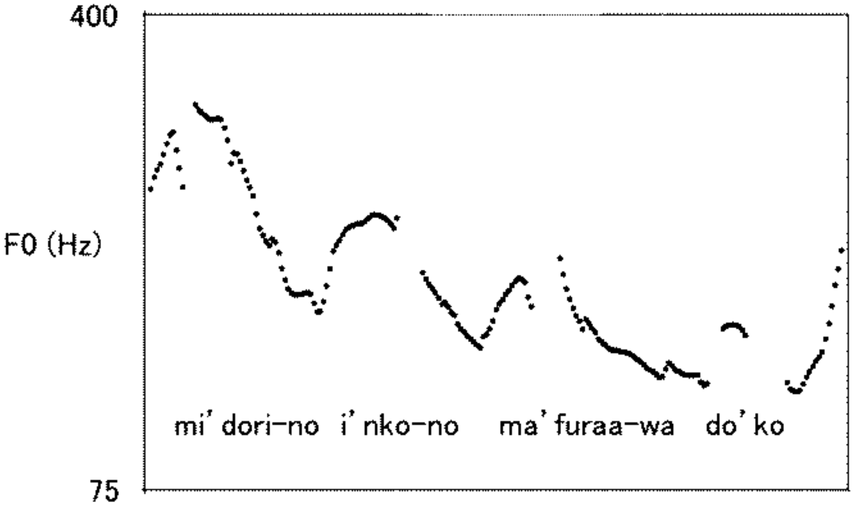



**Figure 2:** F0 contour of the utterance "*mi'dori no i'nko no ma'furaa wa doko?*" with a left-branching structure in which downstep occurs over the three elements. (adapted from Hirose, Arai and Ito 2012)


**Figure 3:** F0 contour of the utterance "*mi'dori no i'nko no ma'furaa wa do'ko?*" with a right-branching structure in which downstep occurs over the three elements.

of a new major phrase are aligned in Japanese. Therefore, (2) actually comprises of two major phrases, with their boundary between *mi'dori no* and *i'nko no* (see also Nagahara 1994; Sugahara 2003).

Kubozono (1988), however, argues that once downstep is triggered by an accented element, it continues irrespective of the branching structure (until the major phrase

is reset for an independent reason). The phenomenon of the elevated pitch peak on the second element is called "metrical boost" because the higher realization of F0 is conditioned by the right-branching structure. Kubozono, following Poser (1984) and Beckman and Pierrehumbert (1986), maintains that metrical boost is driven by the syntactic structure (i.e., branching structure) and the phonological structure (accent status of the constituent), and at the same time, it is a phonetic realization rule that is not directly driven from the syntactic representation, but is mediated by the right- or left-branching prosodic representation within the domain of downstep. According to such a view, the prosodic phrasing status at the major phrase level associated with the two distinct syntactic branching structures is therefore the same if downstep occurs over the three elements irrespective of the syntactic branching structures. Kubozono (1989) proposed that the phonetic realization rule parses the distinct hierarchical structures *within* the domain of downstep, wherein a minor phrase representation can be formed by a binary recursive mechanism. As a result, (1) could have the two possible structures shown in (3), depending on the syntactic branching structure associated with it. Metrical boost takes place in response to a right-branching structure in the minor phrase representation, which reflects the syntactic branching structure. In a production study intended to reconfirm Kubozono (1988), Venditti (1994) reports mixed results, in which the evidence for occurrence/reset of downstep on the second element in the right-branching structure varied among speakers. Although the phonological status of metrical boost is under debate, the elevation of the F0 peak on the second element in a right-branching structure contrasting with the left-branching structure is a widely recognized phenomenon among researchers.

(3) a. $\{_{\text{MajP}}\ \{_{\text{minP}}\ \{_{\text{minP}}$ *mi'dori no* } $\{_{\text{minP}}$ *i'nko no* } } { $_{\text{minP}}$ *ma'furaa* }}
(left branching)

b. $\{_{\text{MajP}}$ { $_{\text{minP}}$ *mi'dori no* } { $_{\text{minP}}\ \{_{\text{minP}}$ *i'nko no* } $\{_{\text{minP}}$ *ma'furaa* } }}
(right branching)

Let us think about the cases where the presence or absence of downstep cannot provide a cue for disambiguation. Since downstep is triggered by a lexically accented element, the prosodic demarcation may become less obvious if the initial modifier does not carry a lexical accent, and hence does not trigger downstep. Consider (4), in which all three elements are lexically unaccented.

(4) *mizuiro no* *kaeru no* *boosi*
blue GEN frog GEN cap

The two branching structures can still be distinguished because the series of [*unaccented modifier* + N1+ N2] tends to be dephrased into one minor phrase in the

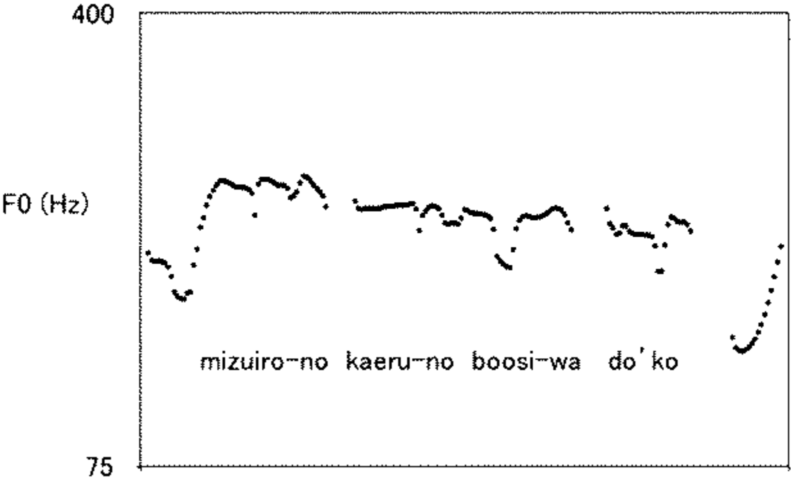



**Figure 4:** F0 contour of the utterance "*mizuiro no kaeru no boosi wa doko?*" with a left-branching structure, in which all three elements are unaccented. (adapted from Hirose, Arai and Ito 2011)

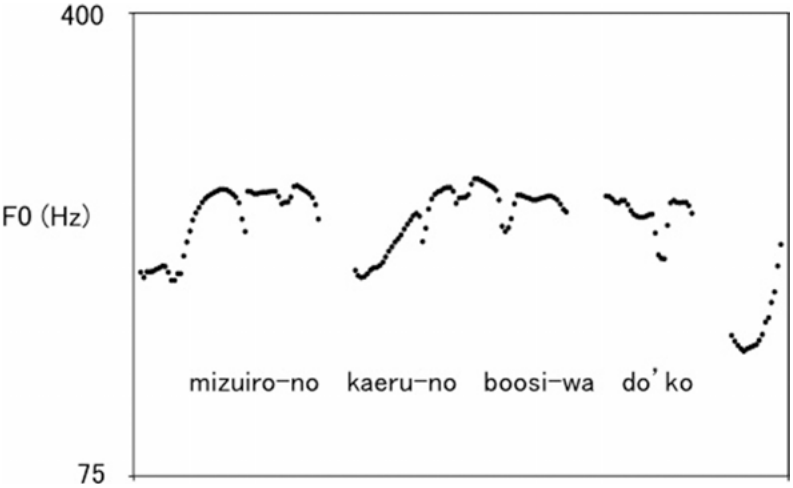



**Figure 5:** F0 contour of the utterance "*mizuiro no kaeru no boosi wa doko?*" with a right-branching structure, in which all three elements are unaccented

left-branching structure (Figure 4). In contrast, the initial delimitative LH sequence (initial rise) at the beginning of the second element marks a minor phrase boundary (L%) in the right-branching structure (Figure 5). Since metrical boost occurs independently from the occurrence of downstep, the pitch range of the second items is also expected to be larger in the right-branching structure. Selkirk, Shinya and Sugahara (2003) argue that the realization of the initial rise at the beginning of a

minor phrase depends on the status of the syntactic boundary (XP vs. nonXP) with which the left edge of the minor phrase is to be aligned. In this chapter we remain open about the syntactic status of the constituent *kaeru no boosi* for now.

So far we have illustrated the prosodic realization of the two branching structures when the three elements are either all accented words, or all unaccented words.

The following sections demonstrate that there are various non-syntactic factors that affect the realization of prosody, reminding us that syntax-prosody correspondence is often not straightforward.

# 3 Processing of the prosodic cues: dealing with an ambiguity in interpreting the prosodic prominence

We have discussed above the observation that a right-branching structure such as (1) and (4) is demarcated by a raised pitch of the second element. On the perception side, if the listeners detect an elevation of the pitch peak (compared to the peak height which would be expected when the downstep would be occurring) on N1, they should process it as a cue indicating the right-branching structure, i.e., the first modifier should be associated with the N2. The scaled judgment study on the edited spoken stimuli by Venditti (1994) varied the relative peak heights between the first and the second items, in addition to the pause duration between the two. The results showed that the difference between the F0 peaks on the two elements correlated with the responses of the listeners forced choice judgment task between the left- and right-branching interpretations. The study further suggested that listeners also exploit other non F0 cues such as the presence of pause to disambiguate between the two interpretations.

In fact, the pitch of a word is also raised if the information carried by that word is emphasized, for example, the word receives focus because the referent of that word stands in a contrastive relationship with some other entity in a given context. Focused elements are associated with a higher F0 peak and an expanded F0 range (Pierrehumbert and Beckman, 1988), followed by a compressed pitch range on the post-focus items. The magnitude of pitch range enlargement is generally greater for accented words (as in (1)) than for unaccented words (as in (4)), but it initiates a new prosodic phrasing (IP) regardless of the accent type (Ito 2002). So, raising of F0 is a common acoustic correlate between metrical boost and focus prominence, although the phonological status between them and the prosodic phenomena applying to the neighboring elements are not the same. This means the function of the elevated pitch on *i'nko no* in (1) or that on *mizuiro no* in (4) could be ambiguous between a cue to syntax (metrical boost signaling a right-branching structure, imposing the meaning 'a green scarf with a parrot'/'a blue cap with a frog') or a focus prominence

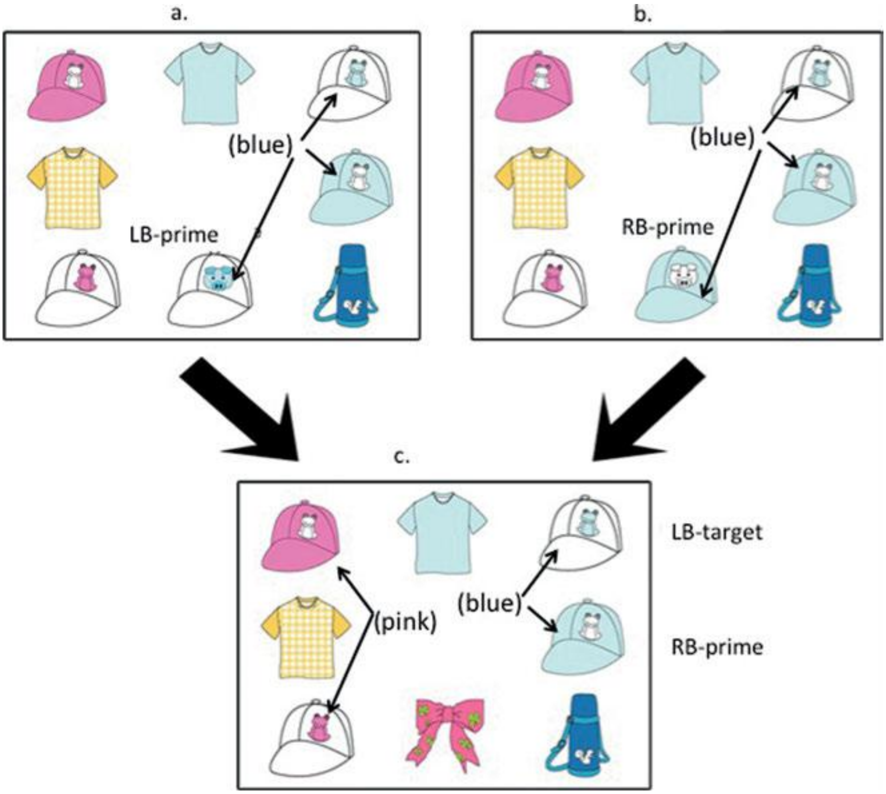



**Figure 6:** Example slides of the LB and RB primes ((a) and (b), respectively) and the target (c) visual display for Experiment 1 (adapted from Ito et al. 2015)

such as "a parrot as opposed to a sparrow"/"a frog as opposed to a lizard" in the real-time processing of the prosodic information.

A series of eye-movement studies using visual world paradigm by Ito, Arai, and Hirose (2015) investigated the native speakers' interpretation of a pitch expansion in referential contexts in which i) the discourse context and the contrastive focus on the second element created felicitous contrastive information and ii) the contrast expressed by the discourse context and the contrastive focus on the second element are not congruent with each other, using the spoken sentences with the left-branching (LB) and right-branching (RB) ambiguity. The experiment consisted of context (prime) – target trial pairs.

In Experiment 1, the subjects listened to spoken stimuli pairs such as "*mizuiro no buta no boosi wa doko?*" (Where's blue pig GEN cap) → "*zyaa*, *mizuiro no kaeru no boosi wa doko?*" (Then, where's blue frog GEN cap), each presented with a visual scene such as (a)–(c) in Figure 6. The target spoken sentence either had a pitch

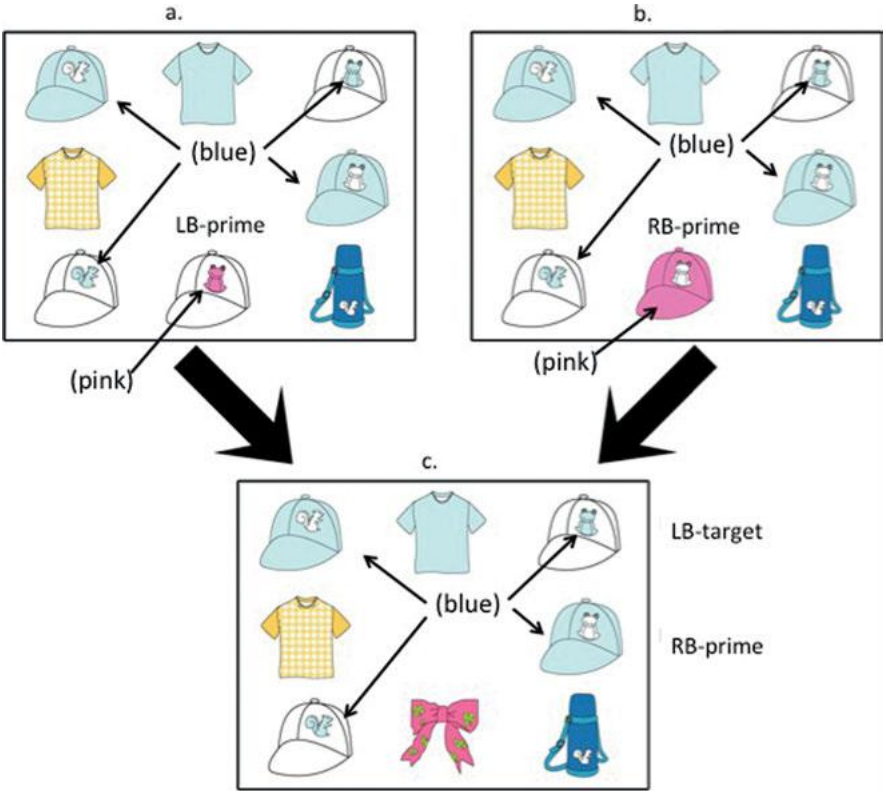



**Figure 7:** Example slides of the LB and RB primes ((a) and (b), respectively) and the target (c) visual display for Experiment 2 (adapted from Ito et al. 2015)

expansion on the second element *kaeru no* or not. The enlarged pitch range (hence higher F0 peak) on the second element could be interpreted as occurrence of metrical boost demarcating an RB structure. Alternatively, the pitch expansion could also be taken to indicate the contrast between the prime (blue pig GEN cap) and the target (blue frog GEN cap).

In Experiment 2, the established contrast between the prime and the target was in fact incongruent with the contrastive focus expressed by prosody. This time, the subjects listened to the spoken stimuli pairs such as "*momoiro no kaeru no boosi wa doko?*" (Where's pink frog GEN cap) → "*zyaa, mizuiro no kaeru no boosi wa doko?*" (Then, where's blue frog GEN cap), each presented with a visual scene such as (a)–(c) in Figure 7.

The target spoken sentence sets were identical with those used in Experiment 1, again with or without the pitch expansion on the second element. This time, however, the contrastive focus interpretation of the pitch expansion would be infelicitous in the referential context, in which the target object in the prime trial and the

two potential target objects in the target trial are both frogs, contrasting in color. Thus, the function of the prosodic prominence on *kaeru* (frog) is restricted to the structural cue.

In both Experiment 1 and Experiment 2, the enlarged pitch range on *kaeru no* obviously counteracted the default bias towards the left-branching interpretation, that is, participants' looks to the LB-target objects were reduced in the condition with the pitch expansion as compared to in the condition without pitch expansion. This presumably means that the pitch expansion was processed primarily as an instance of metrical boost. Interestingly, when the contrastive interpretation of the prosodic manipulation could not possibly signal the contrastive relationship established between the prime and the target, as designed in Experiment 2, the size of such effect (increased looks to the RB-target relative to the LB-target) became larger. The study concludes that processing of prosodic prominence that potentially allows two functions is affected by the visually-provided referential context. An enlarged pitch range/ higher peak f0 on the second element was more likely to be processed as a RB structure-marking cue when the plausibility of contrastive interpretation of such cue is reasonably low. The relationship between syntactic and prosodic representations is not always one-to-one: this study demonstrates that listeners can flexibly deduce the best use of the prosodic cue based on the given context.

This study emphasizes the perceptual similarity between the F0-affecting phenomena of metrical boost, and prominence driven by contrastive focus. However, the acoustic profiles between the two should differ in other aspects. For example, focus prosody is realized not only by the expanded pitch range on the focused element, but also by the post-focal de-accenting phenomena (Deguchi and Kitagawa 2002; Hirotani 2005; Ishihara 2002, 2003, 2004, 2007; Kitagawa 2005; Sugahara 2002, 2003, 2005). Further investigation is still needed to investigate specifically how sensitive listeners are to the post-focal phenomena, to distinguish focus prosody from the pitch boosting phenomena motivated by syntactic or some other factors.

## 4 Production of the prosodic cues: Does speakers' awareness of ambiguity make a difference?

Our next question was to see whether native speakers encode the structural distinction in the distinct prosodic patterns discussed above in situations other than a "please-read-aloud-from-text" type of setting. Hirose (2006) set up an experimental task in which the participants were supposed to give verbal instructions to a confederate (the experimenter). Fourteen speakers, tested one by one, were told to refer to the target object out of four objects simultaneously appearing in a visual display, using the fixed phrase (modifier + N1 + N2) as in (5), so that the confederate could correctly identify the object being mentioned.

(5) *ao'i* *siidi'i no* *ke'esu*
blue CD GEN case

There were two sessions in separate blocks: In the first block, the target visual object was always presented in an "unambiguous context" in which there was only one possible target object, corresponding either to the left- or right- branching interpretation. That was, there was a plastic CD-case with a blue-colored CD in it, together with three other unrelated objects on the same scene, or there was a blue-colored case with a CD in some other color, with three other unrelated objects. In the second block the target visual object was always presented in an "ambiguous context". That is to say, there were two possible targets corresponding to the left- and the right-branching interpretations (described above) on the same display, with two other unrelated objects. In each trial, the participants were presented with the visual scene, with another picture on the side illustrating just the target object, and then were told to refer to the target object on the display using the designated fixed phrase. Their task was to instruct the confederate, who would be presented with the same visual scene, so that s/he could identify the target referent. Figure 8 shows pitch tracks for a pair of representative examples of the collected utterances taken in the "ambiguous context" when the speaker was presumably under some pressure to produce the structurally-ambiguous sentence in an unambiguous manner.

The F0 analysis of the speakers' utterances (i.e., (5)) showed the peak F0 on the first element (*ao'i* 'blue') and that on the second element (N1, i.e. *siidi'i no* 'CD GEN') produced a significantly larger declination for left-branching structures compared to right-branching structures in both ambiguous and unambiguous conditions, confirming the observations in the previous literature. This difference is indicated by the first arrow in Figure 8. The difference was observed regardless of the referential ambiguity in the visual scene (i.e., regardless of whether the participants performed the task in the ambiguous or in the unambiguous context), although the difference was more pronounced in the ambiguous context.

In addition to the predicted difference, we found an unexpected significant difference between the F0 peak heights of the two different branching structures, in both ambiguous and unambiguous conditions. Namely, F0 peaks for the third element were higher than those of the second element in the left branching structure, as illustrated by the second arrow in Figure 8. By contrast, right branching structures showed a constant decline between the second and the third elements compared to left branching.

Do speakers disambiguate the branching structures differently between the ambiguous and the unambiguous contexts? Prosodic correlates reflecting the speakers' attempt to disambiguate the branching structure in response to the experimental manipulation showed up in the duration of the phrase-final segments (i.e., "*i*" as

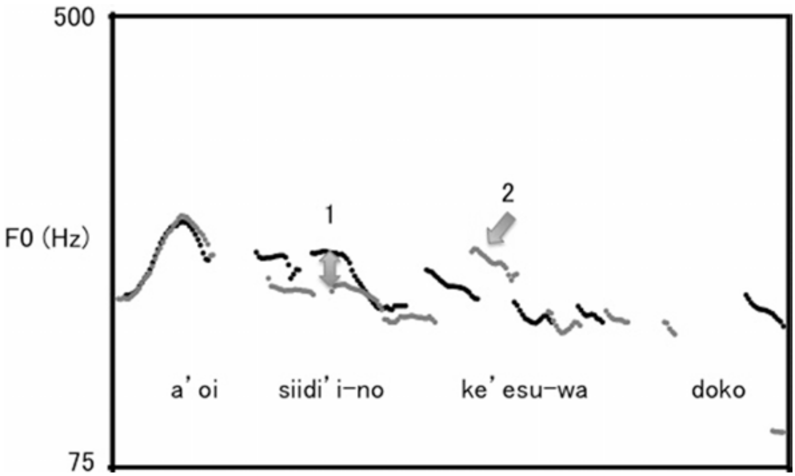



**Figure 8:** Example F0 contours of the utterance "ao'i siidi'i no ke'esu wa doko?" with a right-branching structure (black line) with a left-branching structure (gray line), in the "ambiguous context", produced by the same speaker

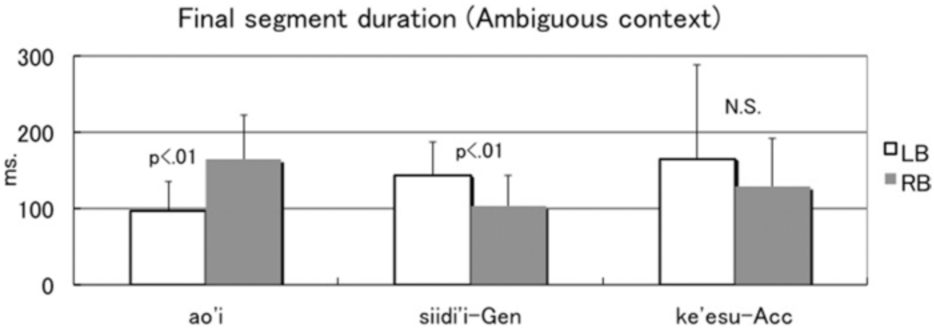



**Figure 9:** The mean phrase-final segment duration (in milliseconds) of the modifier, N1 and N2, for *ao'i siidi'i no ke'esu o* in the ambiguous block, collapsed over 14 speakers

in *ao'i*, "*o*" as in *siidi'i no* or "*o*" as in *ke'esu o*), as shown in Figure 9 (ambiguous context) and Figure 10 (unambiguous context). Interestingly, only in the ambiguous context, was the phrase-final segment of the first element significantly longer when the right-branching was intended compared to when the left-branching was intended. In addition, the phrase-final segment of the second element was longer when the left-branching was intended compared to when the right-branching was intended. In the unambiguous context, no such durational difference was found.

This finding suggests that the reflection of branching structure in the F0 pattern in a (quasi-) communicative situations is at least consistent with the findings in the

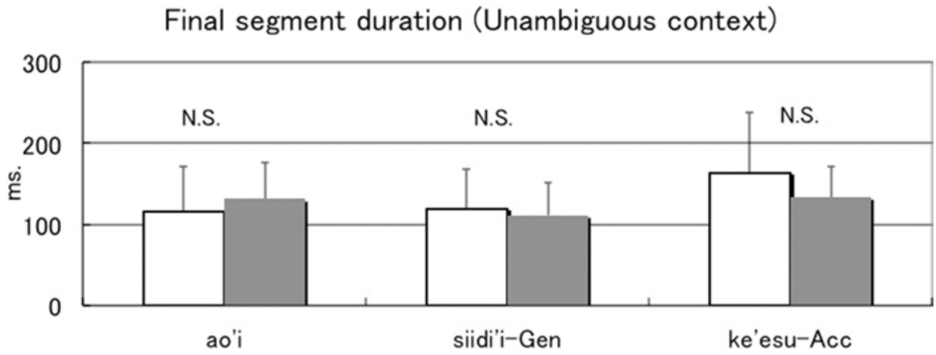



**Figure 10:** The mean phrase-final segment duration (in milliseconds) of the modifier, N1 and N2, for *ao'i siidi'i no ke'esu o* in the unambiguous block, collapsed over 14 speakers

previous literature that come from more formal laboratory settings, except that the relative F0 peak heights between the second element (N1) and the third element (N2) also showed distinct patterns depending on the branching structure. However, when the speakers were aware of the structural ambiguity and particularly wished to produce the phrase in an unambiguous way, they instead resorted to the durational cue to indicate the grouping of the constituents.

## 5 The effect of constituent length and implicit prosody

Besides syntactic structure, there are also extra-syntactic factors that are known to affect the prosodic structure of the modifier + N1-gen N2. Kubozono (1989; 1993) observes that the F0 peak of a predicted downstepping N1 in the left-branching structure is realized higher when the whole phrase involves four items instead of three, as in Figure 11. This was reconfirmed in later studies such as Shinya, Selkirk and Kawahara (2004). This phenomenon is called *rhythmic boost*. In such situations, the difference in the F0 contour for the right-branching and the left-branching structure becomes neutralized to (6), because in both cases the F0 peak on the N1 would be realized higher than the value expected by downstep (by metrical boost in the former and by rhythmic boost in the latter). The elements A and B together and C and D together are dominated by a higher level of minP (so-called superordinate minor phrase, sMiP by Shinya et al. (2004)), by assuming a recursivity in minor phrase.

(6) $\{_{\text{MajP}}\{_{\text{minP}}\{_{\text{minP}}\ A\}\{\{_{\text{minP}}\ B\}\}\ \{\{_{\text{minP}}\ \{\{_{\text{minP}}\ C\}\{\{_{\text{minP}}\ D\}\}\}$

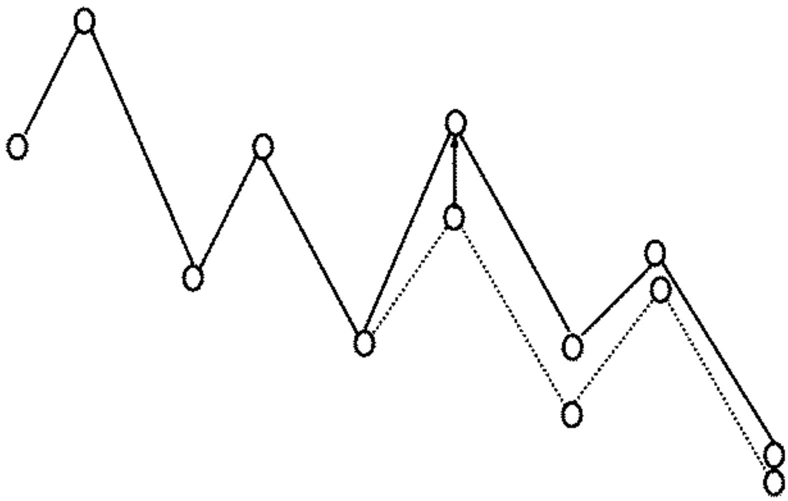


**Figure 11:** Basic downstep contour (solid line) and surface F0 contour (Rhythmic Boost added; dotted line), duplicated from Kubozono (1989: 53)

An alternative view is to regard this phenomenon of pitch elevation as an indication of re-setting the domain of downstep. Selkirk (2000) argues for universal constraints on the minimum and maximum size of prosodic constituents within the framework of Optimality Theory, in addition to constraints sensitive to aspects of syntactic structure. According to this size constraint, the optimal number of minor phrases to constitute a major phrase is two, so the phrase is parsed into two major phrases and therefore the pitch range of the downstep within a major phrase resets at the third element (N1), as schematically shown in (7).

(7) $\{_{\text{MajP}} \{_{\text{minP}} A\}\{_{\text{minP}} B\}\}\{_{\text{MajP}} \{_{\text{minP}} C\}\{_{\text{minP}} D\}\}$

Regardless of whether it reflects a boost within a major phrase or it marks a new major phrase, the fact that the pitch rises on the third element in a left-branching structure leads to a prediction that the preferred interpretation of an NP with a branching ambiguity should modulate with whether the modifier consists of one prosodic word or two (more specifically, one minor phrase or two).

This provides yet another example of "ambiguity" of the prosodic cue, in this case, when the third element out of four accented minor phrases (when rhythmic boost is expected to occur) coincides with the potential right-branching position of the phrase. Consider the two cases of the branching ambiguity in (8) and (9) below.

| | | | |
|---|---|---|---|
| (8) | *hinnona'i* | *suna'kku no* | *ho'sutesu o* |
| | coarse | bar GEN | hostess ACC |

| | | | | |
|---|---|---|---|---|
| (9) | *nantona'ku* | *hinnona'i* | *suna'kku no* | *ho'sutesu o* |
| | somewhat | coarse | bar GEN | hostess ACC |

The modifier phrases considered here are *hinnona'i* 'coarse/ unsophisticated' in (8) and *nantona'ku hinnona'i* 'somewhat coarse/unsophisticated' in (9). In production, (8) should exhibit distinct prosodic patterns depending on the syntactic branching structure intended by the speaker. If the modifier becomes two-phrases long as in (9), the difference in the F0 contour between the two branching structures is expected to become less distinguishable; these both match the right-branching structure where the F0 rises on *suna'kku no*.

There haven't been studies experimentally testing how listeners cope with such ambiguity in the role of the prosodic cue (metrical boost vs. rhythmic boost) during the processing of spoken sentences. What about processing those phrases in reading? It has been proposed that even in silent reading – where the input carries no prosody or punctuation – prosodic contours are created (mentally computed) by the reader and the syntactic processor is sensitive to these (Bader 1998; Fodor 1998, 2002). Fodor (1998) argued that some of these cross-linguistic parsing preferences are of prosodic origin, even in silent reading. The idea was formalized as The Implicit Prosody Hypothesis:

> "The Implicit Prosody Hypothesis (IPH): In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for the construction." (Fodor 2002)

If this were the case for the processing of phrases with the branching ambiguity in silent reading, the above prediction should hold. That is, the branching ambiguity in (8) above should be more likely to be resolved in favor of the right-branching structure compared to (9), since the reader would be projecting the two + two symmetric structure onto the input for (9) while (8) will be assigned the unmarked prosodic contour without the metrical boost, i.e. the prosodic structure corresponding to the left-branching syntactic structure. Fodor and Inoue (1994) in fact report their intuitive judgment supporting this prediction.

A self-paced reading time study reported in Hirose (1999) examined this length effect on the interpretation of the modifier-modificant relationship in an experimental setting. Reading time data were collected for four types of Japanese sentences, illustrated in (10), that were presented frame-by-frame.

(10) a. short AdjP, LB interpretation only (Forced N1 modification):

| | | | |
|---|---|---|---|
| *nisyuukan ma'e* | *usugura'i* | *suna'kku no* | *ho'sutesu o* |
| two weeks ago | dim | bar GEN | hostess-Acc |
| *Satoru ga* | *buzyokusita* | | |
| Satoru NOM | insulted | | |

"Two weeks ago Satoru insulted the hostess of the dimly-lit bar."

b. short AdjP, ambiguous between LB and RB interpretations:

| | | | |
|---|---|---|---|
| *nisyuukan ma'e* | *hinnona'i* | *suna'kku no* | *ho'sutesu o* |
| two weeks ago | coarse | bar GEN | hostess ACC |

| | |
|---|---|
| *Satoru ga* | *buzyokusita* |
| Satoru NOM | insulted |

"Two weeks ago Satoru insulted the hostess of the coarse bar." (LB)
"Two weeks ago Satoru insulted the coarse hostess of the bar." (RB)

c. long AdjP, LB interpretation only(Forced N1 modification):

| | | | |
|---|---|---|---|
| *nantona'ku* | *usugura'i* | *suna'kku no* | *ho'sutesu o* |
| somewhat | dim | bar GEN | hostess ACC |

| | |
|---|---|
| *Satoru ga* | *buzyokusita* |
| Satoru NOM | insulted |

"Satoru insulted the hostess of the somewhat dimly-lit bar."

d. long AdjP, ambiguous between LB and RB interpretations:

| | | | |
|---|---|---|---|
| *nantona'ku* | *hinnona'i* | *suna'kku no* | *ho'sutesu o* |
| somewhat | coarse | bar GEN | hostess ACC |

| | |
|---|---|
| *Satoru ga* | *buzyokusita* |
| Satoru NOM | insulted |

"Satoru insulted the hostess of the somewhat coarse bar." (LB)
"Satoru insulted the somewhat coarse hostess of the bar." (RB)

The critical region in this experiment was N2-ACC, as it should reveal whether the readers had anticipated the N2 to be the modificant of the AdjP. If they did, there should be an increase in reading time reflecting the semantic mismatch between the AdjP and N2 in the Forced N1 modification conditions (10a) and (10c) compared to the ambiguous counterparts ((10b) and (10d), respectively). As can be seen in the mean reading times at the N2-ACC region plotted in Figure 12, there was a significant interaction between the length of AdjP and the AdjP association type. A subanalysis at the N2-ACC region comparing the two AdjP association types for long AdjP sentences yielded a highly significant difference in reading time. This is the predicted garden path effect, and reflects a cost for sentences in which the AdjP is compatible only with N1 and is not a possible modifier of N2. For the short AdjP conditions, by contrast, comparison of the two attachment types did not yield any significant difference. Overall, the results were largely consistent with the idea that long modifiers are more likely to be associated with N2. The evidence of N2-association of the AdjP was present in this experiment only when AdjPs were made up of two-word phrases. This confirms the informal observation of an NP-length effect by Inoue and Fodor (1995), and possibly provides support for the explanation in terms of implicit prosody proposed by Fodor (1998) based on Kubozono (1988).

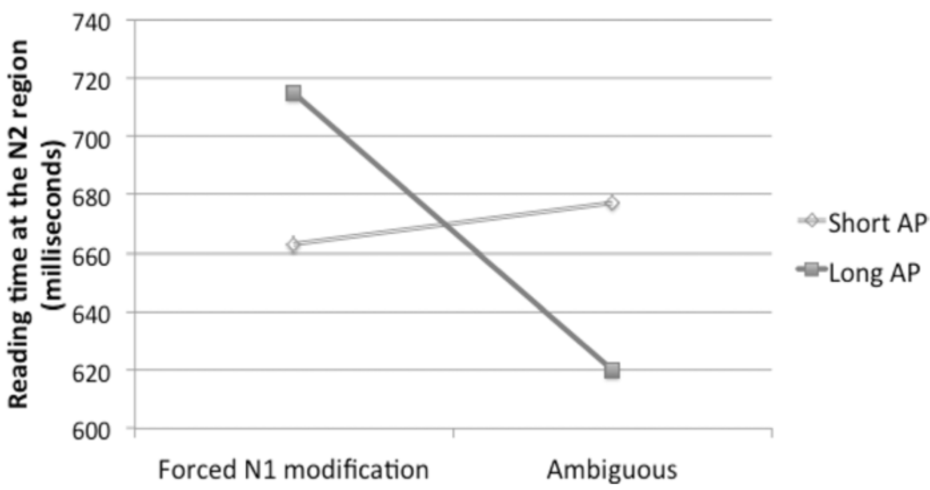



**Figure 12:** The mean reading times (in millisecond) for the N2-ACC region (Region 4)

The above study suggests that prosody can affect sentence processing, even during silent reading. Studies supporting Implicit Prosody Hypothesis, demonstrating length effect on resolving a clause boundary ambiguity can be found in Hirose (2003) and Sato, Kobayashi and Miyamoto (2007).

# 6 Further issues

With respect to the processing bias between the left- and right branching structures in children, very little is known to date. Mazuka and Uetsuki (2004) conducted a study on 4, 5, and 6 year old children together with adult Japanese speakers to see whether the effect of prosody in disambiguation of the branching structure differs among different age groups. In their study, the prosodic difference between the two branching structures was somewhat exaggerated. There was an approximately 700ms pause between the first and the second constituent in the RB-reading, with the same size of pause being placed between the second and the third elements in the LB-reading, in addition to the difference in the F0 information as described above. In the comprehension experiment, the participants were instructed to select a picture consistent with the spoken sentence out of a choice of four pictures. When the spoken sentence accompanied the LB-prosody, adults picked the correct picture in 100% of the cases. Children in all age groups (4, 5, and 6 years old) were fairly successful in selecting the correct picture (73%, 87%, and 83%, respectively, with no statistically reliable difference between age groups). Surprisingly, however, for sentences with the RB- prosody (exaggerated with a huge pause following the modifier), the choice of the correct RB-denoting picture was about the chance level for all

age groups including adults. The overall pattern of the data in adults and children so far does not contradict with the other findings discussed in this section. Provided that the participants rarely selected a non-relevant distracter picture in any trial (Mazuka, p.c.), the percentage at which the participants selected an RB-corresponding picture even with the LB-prosody must have been close to 27%, 13%, and 17% for 4, 5, and 6 years old children, respectively. If so, the data appear to indicate that 4 years olds, the youngest group tested, are fairly willing to assign the RB structure to the ambiguous input. This is still at the stage of speculation and needs to be tested in future studies. Much work needs to be done to look at the developmental aspect of resolution of syntactic ambiguities, including branching ambiguity currently in discussion, and the role of prosody in it.

So far we have limited our discussions to the case of Tokyo Japanese (so-called standard Japanese), but there are various dialects in Japan, exhibiting differences in accentuation at the lexical level and in the details of the tone assignment rules. Besides variations at the lexical level among dialects, is the difference between the two branching syntactic structures projected to different prosodic representations in the same way as in Tokyo Japanese discussed above? Igarashi (2010a) reports that right-branching structure is characterized by the occurrence of metrical boost in Fukuoka Japanese (Igarashi 2007a) and Goshogawara Japanese (Igarashi 2007b) as well in Kumamoto Japanese (Maekawa 1997; Kori 2006a). Studies disagree with respect to whether that is true for Osaka dialect: Sugito (2001) reports that the branching structure is not reflected in the F0 contour whereas Kori (1989; 2006b) maintains that the right-branching structure accompanies a higher F0 on the second element in Osaka Japanese, although the size of the difference between the two branching structures is smaller than that for Tokyo Japanese. Igarashi (2010a) reports results of a production study of six Osaka speakers which support Kori, but also points out that the difference is less likely to show up on certain combinations of accent types (Igarashi 2010b, 2014). To date, the data size is still too limited to make generalizations about the relationship between the syntactic and the prosodic representations in Osaka Japanese. Further research is needed to accumulate empirical evidence in different dialects in Japan, especially in Osaka Japanese.

## 7 Summary

As we have seen so far, the role of prosody in production and comprehension of NPs with a branching ambiguity is less clear than suggested by the theoretical literature. The relationship between syntactic structure, prosodic structure, and prosodic realization depends on a number of factors. The two factors described in this chapter are a phonological factor (lexical accent and phonological size of the constituents) and a discourse factor (the referential context that the listeners are presented with and speakers' ambiguity awareness evoked from the visual context).

More comprehensive research is necessary to determine exactly how speakers choose among various different types of cues (especially those that are not documented in the theoretical literature discussing the syntax and prosody correspondence) to convey their intended meaning under different situations. Also of further interest is what type of cues, other than occurrence of metrical boost, the listeners focus on in comprehension, in various situations which require various degrees of ambiguity awareness.

## References

Allbritton, David, Gail McKoon, and Roger Ratcliff. 1996. Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory and Cognition* 22. 714–735.

Bader, Markus. 1998. Prosodic influences on reading syntactically ambiguous sentences. In Janet Dean Fodor and Fernanda Ferreira (eds.), *Reanalysis in sentence processing*, 1–46. Dordrecht: Kluwer Academic Publisher.

Beckman, Mary E. 1996. The parsing of prosody. *Language and Cognitive Processes* 11(1/2). 17–67.

Beckman, Mary E. and Janet B. Pierrehumbert. 1986. Intonational structure in English and Japanese. *Phonology Yearbook* 3. 255–309.

Cutler, Anne, Delphine Dahan and Wilma van Donselaar. 1997. Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40(2). 141–201.

Cutler, Anne and Takeshi Otake. 1996. The processing of word prosody in Japanese. In Paul McCormack and Alison Russell (eds.), *Proceedings of the 6th Australian International Conference on Speech Science and Technology*, 599–604. Canberra: Australian Speech Science and Technology Association.

Deguchi, Masanori and Yoshihisa Kitagawa. 2002. Prosody and wh-questions. In Masako Hirotani (ed.), *Proceedings of the 32nd annual meeting of the North East Linguistics Society*, 73–92. Amherst MA: GLSA, University of Massachusetts.

Fodor, Janet Dean. 1998. Learning to parse? *Journal of Psycholinguistic Research* 27(2). 285–319.

Fodor, Janet Dean. 2002. Psycholinguistics cannot escape prosody. Paper presented at the Speech Prosody International Conference, 83–90. Aix-en-Provence, France.

Fodor, Janet Dean and Atsu Inoue. 1994. The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research* 23(5). 407–434.

Haraguchi, Shosuke. 1977. *The tone pattern of Japanese: An autosegmental theory of tonology*. Tokyo: Kaitakusha.

Hirose, Yuki. 1999. *Resolving reanalysis ambiguity in Japanese relative clauses*. New York: The City University of New York dissertation.

Hirose, Yuki. 2003. Recycling prosodic boundaries. *Journal of Psycholinguistic Research* 32(2). 162–195

Hirose, Yuki. 2006. Missed cues: Speaker-hearer mismatch and variability. Paper presented at The Nineteenth Annual CUNY Conference of Human Sentence Processing, The City University of New York.

Hirose, Yuki, Manabu Arai and Kiwako Ito. 2011. Influence of contrastive prosody on structural priming. Paper presented at MAPLL (Mental Architecture for Processing and Learning of Language) 2011, Hiroshima, Japan.

Hirose, Yuki, Manabu Arai and Kiwako Ito. 2012. What is in contrast? – The role of prosodic prominence in ambiguity resolution. Paper presented at MAPLL (Mental Architecture for Processing and Learning of Language) 2012, Yamagata, Japan.

Hirotani, Masako. 2005. *Prosody and LF interpretation: Processing Japanese wh-questions*. Amherst, MA: University of Massachusetts dissertation.

Igarashi. Yosuke. 2007a. Pitch range compression and pitch accent deletion in Fukuoka Japanese. *Proceedings of the 21st general meeting of the Phonetic Society of Japan*, 111–116. Nagoya, Japan.

Igarashi, Yosuke. 2007b. Pitch pattern alternation in Goshogawara Japanese: Evidence for a prosodic phrase above the domain for downstep. *Proceedings of the Interspeech 2007*, 434–437. Antwerp, Belgium.

Igarashi, Yosuke. 2010a. Tōgoron ni okeru edawakare kōzō wa donoyōni han'ei sareru no ka – Kinki hōgen to Tōkyō hōgen no baai [How a branching structure is projected in syntax in case of Kinki and Tokyo dialects]. Presented at The 24th Annual Convention of The Phonetic Society of Japan. Kokugakuin University, Tokyo, Japan.

Igarashi, Yosuke. 2010b. Kinki hōgen ni okeru tōgo kōzō to inritu kōzō no kankei wa dō natte irunoka [What is the relationship between syntactic and prosodic structures in Kinki dialect?]. Paper presented at the 321st regular meeting of Phonetic Society of Japan. National Institute for Japanese Language and Linguistics, Tokyo.

Igarashi, Yosuke. 2014. Typology of intonational phrasing in Japanese dialects. In Sun-Ah Jun (ed.), *Prosodic typology*, the second volume, 464–492. New York: Oxford University Press.

Igarashi, Yosuke, Hideaki Kikuchi and Kikuo Maekawa. 2007. Inritsu jōhō [Prosodic information]. In Kikuo Maekawa (ed.), *Nihongo hanashikotoba kōpasu no kōchikuhō* [Construction of the corpus of spontaneous Japanese], 347–453. Tokyo: National Institute for Japanese Language.

Inoue, Atsu and Janet Dean Fodor. 1995. Information-paced parsing of Japanese. In Reiko Mazuka and Noriko Nagai (eds.), *Japanese sentence processing*, 9–63. Hillsdale, NJ: Lawrence Erlbaum.

Ishihara, Shinichiro. 2002. Invisible but audible wh-scope marking. In Line Mikkelson and Christopher Potts (eds.), *Proceedings of the 21st West Coast Conference on Formal Linguistics*, 180–193. Somerville, MA: Cascadilla Press.

Ishihara, Shinichiro. 2003. *Intonation and interface condition*. Cambridge, MA: Massachusetts Institute for Technology dissertation.

Ishihara, Shinichiro. 2004. Prosody by phase: Evidence from focus intonation-wh scope correspondence in Japanese. In Shinichiro Ishihara, Michaela Schmitz and Anne Schwarz (eds.), *Interdisciplinary studies on information structure 1: Working Papers of SFB632*, 77–119. Potsdam: University of Potsdam.

Ishihara, Shinichiro. 2007. Major phrase, focus intonation, multiple Spell-Out. *The Linguistic Review* 24. 137–167.

Ito, Kiwako. 2002. *The interaction of focus and lexical pitch accent in speech production and dialogue comprehension: Evidence from Japanese and Basque*. Urbana-Champaign, IL: University of Illinois dissertation.

Ito, Kiwako, Manabu Arai and Yuki Hirose. 2013. The interpretation of phrase-medial prosodic prominence in Japanese: Is it sensitive to visual and discourse context? *Language, Cognition and Neuroscience* 30(1–2). 167–196.

Kahraman, Barış and Hiromu Sakai. 2015. Relative clause processing in Japanese: Psycholinguistic investigation into typological differences. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Kamide, Yuki and Don C. Mitchell. 1997. Relative clause attachment: Nondeterminism in Japanese parsing. *Journal of Psycholinguistic Research* 26(2). 247–254.

Kitagawa, Yoshihisa. 2005. Prosody, syntax and pragmatics of wh-questions in Japanese. *English Linguistics* 22. 302–346.

Kitagawa, Yoshihisa and Yuki Hirose. 2012. Appeals to prosody in Japanese wh-interrogatives – Speakers' versus listeners' strategies. *Lingua* 122(6). 608–641.

Kraljic, Tanya and Susan E. Brennan. 2005. Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology* 50. 194–231.

Kori, Shiro. 1989. Hatsuwa no onchō o kitei suru yōin – Nihongo intonēshonron – [Decisive factors for intonation – Theory of Japanese intonation]. *Yoshizawa Norio Kyōju tsuitō ronbunshū* [Professor Norio Yoshizawa memorial papers], 116–127. Tokyo: Tokyo University of Foreign Studies.

Kori, Shiro. 2006a. Kumamotoshi oyobi shūhen no hiteikei akusento hōgen ni okeru goonchō to onchōku no keisei [Lexical intonation and intonation phrase formation in expressions with negative accent in Kumamoto City and its vicinity]. *Journal of the Phonetic Society of Japan* 10(2). 43–60.

Kori, Shiro. 2006b. Inritsu tokuchō no chiikisa [Regional differences in characteristics of intonation]. In Keikichi Hirose (ed.), *Inritsu to onseigengo jōhōshori [Information processing of intonation and spoken language]*, 50–64. Tokyo: Maruzen,

Kubozono, Haruo. 1988. *The organization of Japanese prosody*. Edinburgh: University of Edinburgh dissertation.

Kubozono, Haruo. 1989. Syntactic and rhythmic effects on downstep in Japanese. *Phonology 6*. 39–67.

Kubozono, Haruo. 1993. *The organization of Japanese prosody*. Tokyo: Kurosio.

Maekawa, Kikuo. 1997. Akusento to intonēshon – akusento no nai chiiki [Accent and intonation – Accentless areas]. In Ryoichi Sato, Shinji Sanada, Masanobu Kato, and Shuichi Itabashi (eds.), *Nihongo onsei 1: Shohōgen no akusento to intonēshon [Japanese sound 1: Accent and intonation in various dialects]*, 97–122. Tokyo: Sanseido.

Maekawa, Kikuo, Hideaki Kikuchi, Yosuke Igarashi and Jennifer Venditti. 2002. X-JToBI: An extended J-ToBI for spontaneous speech. *Proceedings of International Conference of Spoken Language Processing*, 1545–1548.

Mazuka, Reiko and Miki Uetsuki. 2004. Children's use of prosody in the comprehension of syntactically ambiguous sentences. Poster presented at the 17th Annual CUNY Conference on Human Sentence Processing, University of Maryland.

McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.

Miyamoto, Edson T., Michiko Nakamura and Shoichi Takahashi. 2004. Processing relative clauses in Japanese with two attachment sites. In Keir Moulton and Matthew Wolf (eds.), *Proceedings of the 34th annual meeting of the North East Linguistics Society*, 441–452.

Nagahara, Hiroyuki. 1994. *Phonological phrasing in Japanese*. Los Angeles, CA: University of California at Los Angeles dissertation.

Otake, Takashi and Anne Cutler. 1997. Early use of pitch accent in Japanese spoken word recognition. *Journal of Acoustical Society of America 102*. 3202.

Pierrehumbert, Janet B. and Mary E. Beckman. 1988. *Japanese tone structure*. Cambridge: MIT Press.

Poser, William John. 1984. *The phonetics and phonology of tone and intonation in Japanese*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Sato, Kaori., Mari Kobayashi and Edson T. Miyamoto,. 2007. Lack of implicit prosody effects in deaf readers of Japanese. *Journal of Japanese Linguistics* 23. 35–46.

Schafer, Amy J., Shari R. Speer and Paul Warren. 2005. Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In John C. Trueswell and Michael K. Tanenhaus (eds.), *Approaches to studying world situated language use:*

*Psycholinguistic, linguistic and computational perspectives on bridging the product and action tradition*, 209–225. Cambridge: MIT Press.
Selkirk, Elisabeth O. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press.
Selkirk, Elizabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook 3*. 371–405.
Selkirk, Elisabeth O. 2000. The interaction of constraints on prosodic phrasing. In Gösta Bruce and Merle Horne (eds.), *Prosody: Theory and experiment: Studies presented to Gösta Bruce* (Text, speech, and information technology 14). 231–261. Dordrecht: Kluwer Academic Publisher.
Selkirk, Elizabeth, Takahito Shinya and Mariko Sugahara. 2003. Degree of initial lowering in Japanese as a reflex of prosodic structure organization. In Maria-Josep Solé, Daniel Recasense and Joaquín Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, 491–494.
Selkirk, Elizabeth and Koichi Tateishi. 1991. Syntax and downstep in Japanese. In Carol Georgpoulous and Roberta Ishihara (eds.), *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, 519–544. Dordrecht: Kluwer Academic Publisher.
Shattuck-Hufnagel, Stefanie and Alice E. Turk. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25(2). 193–247.
Shinya, Takahito, Elisabeth Selkirk and Shigeto Kawahara. 2004. Rhythmic boost and recursive minor phrase in Japanese. *Proceedings of the Second International Conference on Speech Prosody*, 183–186.
Snedeker, Jesse and John Tureswell. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language* 48. 103–130.
Speer, Shari and Allison Blodgett. 2006. Prosody. In Matthew J. Traxler and Morton Ann Gernsbacher (eds.), *Handbook of psycholinguistics*, 505–537. 2nd edition. Amsterdam: Elsevier.
Sugahara, Mariko. 2002. Conditions on post-Focus dephrasing in Tokyo Japanese. In Bernard Bel and Isabelle Marlien (eds.), *First International Conference on Speech Prosody*, 655–659. Aix-en-Provence, France.
Sugahara, Mariko. 2003. Downtrends and post-FOCUS intonation in Japanese. Amherst, MA: University of Massachusetts dissertation.
Sugahara, Mariko. 2005. Post-FOCUS prosodic boundaries in Tokyo Japanese: Asymmetric behavior of an F0 cue and domain-final lengthening. *Studia Linguistica* 59. 144–173.
Sugito, Miyoko. 2001. Bunpō to nihongo no akusento oyobi intonēshon – Tokyō to Ōsaka no baai [Grammar, Japanese accent and intonation: Tokyo and Osaka]. In Spoken Language Working Group (ed.), *Bunpō to onsei [Grammar and sound]*, 197–210. Tokyo: Kurosio.
Sugiyama, Yukiko. 2012. *The production and perception of Japanese pitch accent.* Newcastle upon Tyne: Cambridge Scholars Publishing.
Venditti, Jennifer J. 1994. The influence of syntax on prosodic structure in Japanese. In Jennifer J. Venditti (ed.), *Papers from the linguistics laboratory. Ohio State University Working Papers in Linguistics*, vol. 44. 191–223.
Venditti, Jennifer J. 2005. The J_ToBI model of Japanese intonation. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, 172–200. Oxford: Oxford University Press.
Venditti, Jennifer J. 2006. Prosody in sentence processing. In Mineharu Nakayama, Reiko Mazuka and Yasuhiro Shirai (eds.), *The handbook of East Asian psycholinguistics, Volume II: Japanese*, 208–217. Cambridge: Cambridge University Press.
Venditti, Jennifer J., Sun-Ah Jun and Mary E. Beckman. 1996. Prosodic cues to syntactic and other linguistic structures in Japanese, Korean and English. In James L. Morgan and Katherine Demuth (eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 287–311. Mahwar: Lawrence Erlbaum Associates, Inc.

Franklin Chang

# 12 The role of learning in theories of English and Japanese sentence processing

## 1 Introduction

Languages like English and Japanese differ in the word orders that they use. The Japanese translation for (1a) could be (1b).

(1) a. *He gave the book to a man.*

b. (*Kare ga*) *otoko ni hon o ageta*.
He NOM man DAT book ACC gave

This translation illustrates the fact that in Japanese, verbs like *ageta* 'gave' are placed at the end of the utterance and pronouns like *kare* 'he' are often omitted, when they are inferable from the situation. Another difference between English and Japanese is that the Japanese utterance uses particles like *ni* and *o* to mark the case of the nouns. For example, the book is the object being given and it is marked with the *o* particle (*hon o*). Case marking of noun phrases allows Japanese speakers to *scramble* phrases while maintaining the same meaning. For instance, (2a) has a similar meaning to (2b) and this alternation approximates the English alternation between (3a) and (3b).

(2) a. *hon o otoko ni ageta.*

b. *otoko ni hon o ageta.*

(3) a. *He gave the book to the man.*

b. *He gave the man the book.*

These language differences show that English and Japanese speakers have learned a range of very different constraints on word orders in their respective languages.

Linguistic theories characterize syntactic knowledge in English and Japanese in terms of abstract syntactic categories and hierarchical tree structures (Chomsky 1957). This approach has allowed these theories to explain differences between these two languages in terms of linguistic parameters that specify how these hierarchical tree structures are built (Chomsky and Lasnik 1993; Culicover 1997). For example,

theories have posited a verb-direction parameter, which specifies the position of the verb within its phrase. When this verb-direction parameter places the verb early, then the English pattern is produced. When the verb-direction parameter places the verb late in the phrase, then the verb-final Japanese pattern is created. Other parameters have been posited which can explain variation in English and Japanese in whether arguments are omitted or not, and the position of relative clauses with respect to the noun phrase that they modify; Japanese relative clauses place the relative clause before the noun that it modifies as in (4a), while English has the opposite pattern as in (4b).

(4) a. *otoko* *ni* *ageta* *hon*
man DAT gave book

b. *the book which he gave to the man*

Critically, parameters allow the same linguistic machinery (e.g., structure-building rules) to work in different languages. This means that most of the theory is universal and the learner must only adjust the parameters to deal with language-specific behavior.

One motivation for positing universal categories and parameters is to help to explain how syntax is learned (Yang 2003). For example, instead of learning the order of each individual verb in Japanese, one can simply associate Japanese verbs with a universal syntactic category of VERB and then learn how that category is sequenced relative to its arguments. Although setting the parameter can be simple, using the parameter requires the linking of language-specific words to universal categories (Pinker 1984) and in many cases, this can be non-trivial (Mazuka 1998). For example, the English verb *clean* is expressed as a multi-word sequence in Japanese *soozi o suru* (掃除をする) with a light verb combined with an accusative case-marked noun. It is not clear if this multi-word sequence be associated with a universal verb category or whether the light verb *suru* is the verb and the *soozi* is a verbal noun with restricted properties (Miyamoto 1999), but this choice has implications for using a verb direction parameter in Japanese as the light verb+verbal noun account requires additional Japanese-specific rules preventing scrambling of the verbal noun.

The universal categories and parameters approach to syntax acquisition assumes that linguistic rules are sufficiently difficult that they could not be learned without some pre-existing knowledge. But speakers must learn similar abstract culture-specific rules in other domains outside of language. For example, Japanese speakers tend to park the head of their cars facing outwards in parking lots (head-last, left side of Figure 1), while English speakers tend to park the head of car facing inward (head-first, right side of Figure 1).

**Figure 1:** Japanese head-last parking (left) and American head-first parking (right)

Like linguistic rules, these cultural parking rules are not just memorized sequences of actions, because they can generalize to novel items. If you rent a van, you will tend park it in the direction that is specified by your cultural rule, even if you have never driven a van before. Assuming that innate head-direction parking parameters are not present, the fact that humans can implicitly learn these rules that apply to abstract categories (e.g., VEHICLE) suggests cross-cultural variation can be explained by domain-general learning mechanisms. Here we apply this idea to language and explore whether a domain-general learning mechanism, in the form of a connectionist model, can explain variation in English and Japanese sentence production.

It is normally assumed that linguistic rules are learned during language acquisition and then universal processing mechanisms makes use of these language-specific rules to parse and produce language (Bock and Levelt 1994; van Gompel and Pickering 2007). But just as non-linguistic behaviors show great variability across cultures, linguistic behaviors and the processes that support them might also differ greatly across languages (Croft 2001; Evans and Levinson 2009). A strong motivation for the idea that language processing differs across languages comes from the fact that processing is incremental. This means that the choices and expectations at different sentence positions are sensitive to language-specific word and structure choices at those positions (Altmann and Kamide 1999; Kamide, Altmann and Haywood, 2003). Since different languages have a diverse set of word and structural choices at different points in sentences, speakers and listeners should naturally have processing biases that reflect the particular language being used. For example, the different position of the verb in English and Korean (which has a verb position that is similar to Japanese) might play a role in explaining differences in eye-tracking in these languages (Choi and Trueswell 2010; Snedeker and Trueswell 2004). Thus if position-specific expectations are used in sentence processing, then incremental language processing must be different in languages with different word orders.

In this chapter, processing will be used to refer to both sentence comprehension and production, as it is thought that similar types of representations are useful for both processes (Vosse and Kempen 2000). To explore the possibility that processing

may be done in language-specific ways, I will first describe a connectionist model that links language learning to sentence production. Then I will review theories that have posited universal mechanisms to explain processing biases in English and Japanese sentence production, and an alternative account will be presented based on the model (Chang 2009). The model learns language-specific syntactic representations and uses these representations in incremental sentence production. It will be argued that the model offers a more parsimonious account of these psycholinguistic phenomena, where processing biases arise out of language-specific knowledge. The implications of this approach for language development will also be discussed.

## 2 An integrated model of acquisition and production in English and Japanese

One approach to learning internal representations is to use connectionist learning algorithms, which are computational learning mechanisms that are inspired by the processes that take place in the brain (Dell, Chang and Griffin 1999). The brain is made up of a network of neurons and learning involves changing the graded strength of the connections between the neurons in this network in response to the environment. Processing in the brain's neural network is thought to involve spreading activation between neurons in ways that reflect the strength of their connections. Connectionist models are computer simulations that make use of spreading activation in a network to account for human behavior. These models use artificial neuron-like units which have weights that represent the strength of their connection to other neurons, and these weights may be learned using different connectionist algorithms. Connectionist models have been applied to phonology (Dell, Juliano and Govindjee 1993), morphology (Plunkett and Juola 1999), lexical knowledge (Oppenheim, Dell and Schwartz 2010), syntax (Elman 1990), and semantics (Rogers and McClelland 2004). There are models of important linguistic phenomena like reading (Ans, Carbonnel and Valdois 1998; Plaut, McClelland, Seidenberg and Patterson 1996), aphasia (Dell, Schwartz, Martin, Saffran and Gagnon 1997), parsing (Palmer-Brown, Tepper and Powell 2002), speech errors (Dell 1986), and recursion (Christiansen and Chater 2001). In particular, these algorithms have been applied to various Japanese phenomena like kanji and phoneme recognition (Dominey, Hoen and Inui 2006; Ijuin et al. 2000; Joe, Mori and Miyake 1990; Mori and Yokosawa 1989; Negishi 2006; Tsuzuki 1996; Waibel, Hanazawa, Hinton, Shikano and Lang, 1988). But importantly, these same learning algorithms have also been applied in non-linguistic domains such as recognizing heart attacks (Baxt and Skora 1996), predicting the stock market (Enke and Thawornwong 2005), and driving cars (Pomerleau 1993).

Therefore, these algorithms are powerful domain-general learning mechanisms that have been successful at learning both linguistic and non-linguistic representations.

One of the main differences between a connectionist approach to language and approaches based on linguistic theories is the tightness of the link between learning and processing. In linguistic parameter theories (Culicover 1997), setting of parameters takes place in development and once they are set, parameters do not change in adults. In connectionist approaches, learning involves gradual adaptation of the representations that support processing in response to the environment. Since these changes are small, this type of learning can continue into adulthood and may be involved in processing phenomena (as we will see later with structural priming). Another difference between these approaches is with respect to the nature of the representations in different languages. Linguistically-oriented processing theories are often defined in terms of their assumptions about levels of representation (e.g., Bock and Levelt (1994) argue for distinct functional and positional levels in sentence production) and it has been argued that similar level-specific mechanisms are used in different languages (Tanaka, Branigan, McLean and Pickering 2011). Connectionist approaches, on the other hand, assume that processing in different languages can be very different, because the internal representations that support processing are learned from the input rather than determined by the theory.

To explore the relationship between learning and processing, we will focus on a connectionist model of language acquisition and sentence production called the Dual-path model (Chang, Dell and Bock 2006; Chang 2002, 2009; Fitz and Chang 2008). The model learns its language representations from message-sentence pairs and it uses these pairs to acquire word-concept links and syntactic constructions. It can learn English- or Japanese-like languages when given appropriate input and it is able to do this without innate linguistic categories or parameters. In this section, the different components of the model will be described and its behavior motivated.

Incremental sentence processing requires the ability to generate expectations about words and word categories at different positions in sentences. One connectionist system that is able to develop these types of expectations is a simple recurrent network (*SRN*, Elman 1990). SRNs learn their internal representations by trying to predict the next word in a sequence and making small changes to their internal representations to make them better at predicting that word in the future. SRNs are composed of a feed-forward architecture where an input layer is linked to a hidden layer and then via this hidden layer to the output layer (bottom half of Figure 2). During training, the input layer is provided with the previous word and the output layer is given the next word as its target. The model learns weights between the input and hidden layers and between the hidden and output layers that help it to predict the next word based on the previous word.

However, SRNs also have an extra layer called the context layer, which contains a copy of the previous hidden layer activation. This context layer is connected to the hidden layer, allowing the SRN to learn longer-distance dependencies, such as

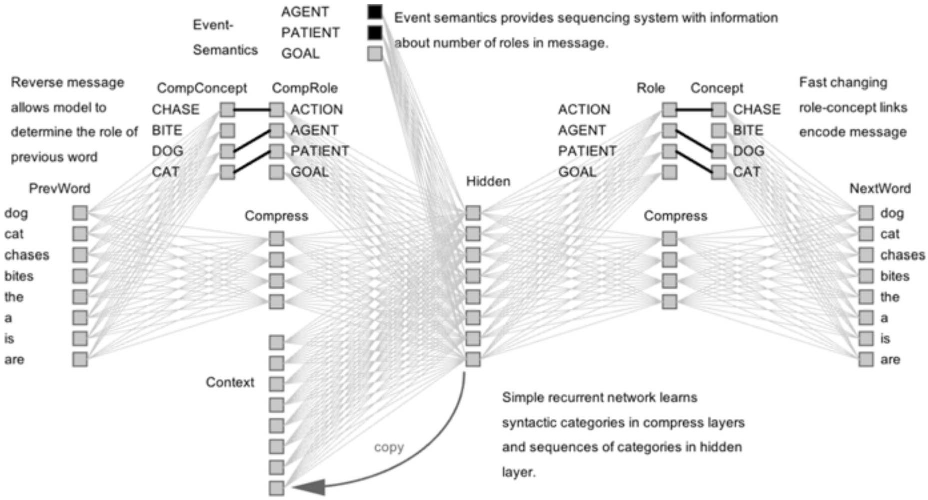



**Figure 2:** English version of the Dual-path model (Chang, Dell and Bock 2006; Chang 2002, 2009; Fitz and Chang 2008)

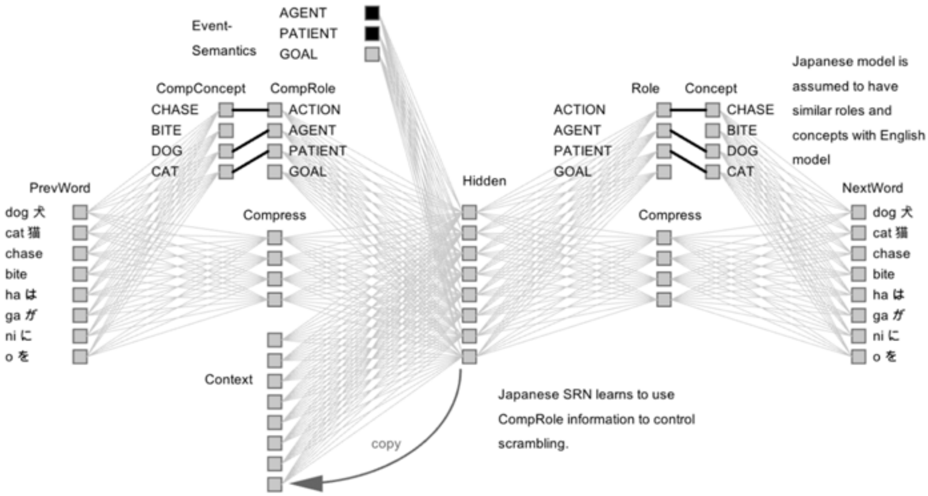



**Figure 3:** Japanese version of the Dual-path model (Chang 2009)

subject-verb agreement across embedded clauses (Christiansen and Chater 2001). Furthermore, when additional compression layers are placed between the input and hidden layers and hidden and output layers (Elman 1993), these layers compress the distinctions that can be encoded, causing the model to learn syntactic categories in these layers. The end result is that the model will develop representations that allow it to activate/predict various sets of words (e.g., nouns, verbs) at each point in sentences. Figures 2 and 3 provide a depiction of the English and Japanese

Dual-path models, respectively; the only difference between the two models at the beginning of training is the input and output word layers. To emphasize the similarity between the English and Japanese models, both models used the same set of content words (e.g., *dog*, *chase*). Thus, the main representational difference in the models was the function words in the English model (e.g., *the*) and the particles in the Japanese model (e.g., *ni*, *o*, which are treated as words). The SRN that instantiated the sequencing pathway can be seen in the bottom half of each model in Figures 2 and 3, where the input layer PrevWord maps through a smaller Compress layer to the Hidden layer and then through a smaller Compress layer to the output layer NextWord (Hidden layer activations are copied back to the Context layer).

The simple recurrent network with compression layers will predict categories of words at each point in a sentence. But sentence production requires that individual words are selected rather than categories, and these incremental word choices should match the message that the speaker intends to convey. For example, the message conveyed by the sentence *the dog chased the cat* is different from the message conveyed by *the cat chased the dog*. It is thought that the difference in meaning is due to how thematic roles like agent and patient are assigned to concepts in these utterances (Fodor and Pylyshyn 1988; Jackendoff 1992). In the first sentence, the dog is the agent of chasing and the cat is the patient, and in the second sentence, the reverse is the case. When message information was given directly to an SRN, it learned message-specific sentence representations that did not generalize (abstract syntax was not learned, Chang 2002). The Dual-path model addressed this problem by using separate pathways for meaning and sequencing regularities.

The meaning pathway in the Dual-path model represented the message as a set of links between roles in a Role layer and concepts in a Concept layer (top right in Figures 2 and 3). For example, for Japanese sentence (5), the agent role was linked to the concept DOG and the patient role was assigned to the concept CAT (for simplicity, roles like agent and patient will be used in this chapter, but the actual model had a slightly different *XYZ* theory of roles, see Chang 2002).

| | | | | | |
|---|---|---|---|---|---|
| (5) | *inu* | *ga* | *neko* | *o* | *oikaketa.* |
| | dog | NOM | cat | ACC | chased |
| | 'The dog chased the cat.' | | | | |

The Concept layer was linked to the model's output word layer (NextWord layer in Figures 2 and 3). Through cross-situational learning (Yu and Smith 2007), the model learned that the DOG concept was associated with the word *inu* and the CAT concept was associated with the word *neko*. Since the message linked roles and concepts and the concepts were also connected to words via learned links, activating a given role unit led to the production of the word for the concepts that were linked to that role.

The Hidden layer of the SRN was also connected to the Role layer, so the model could learn how to activate role units in order to facilitate prediction at particular sentence positions. For example after the phrase *inu ga*, the SRN learned to activate the patient role unit in the Role layer, which then activated the concept CAT, which helped to predict the actual next word as *neko*. The Hidden-Role-Concept-NextWord part of the meaning pathway therefore allowed the model to learn language-specific ways to activate roles at different sentence positions.

The role-concept links in the message were also used to signal the use of pronouns and argument omission (sometimes viewed as pronouns that lack phonology). Sentences in English with pronouns like *he chased the cat* were translated with argument omission in Japanese as *neko o oikaketa*. It is assumed that pronouns and argument omission are used when the listener is able to recover the referent from the discourse, the situation, or world knowledge (e.g., *he* refers to some male in the world). This was implemented in the model by reducing the strength of the link between the role and the concept to 50% of its normal strength, because the speaker knows that the listener does not need a full description in order to retrieve the concept. Reducing the strength of the concept leads to argument omission, but it does not explain why some languages like English require pronouns to be produced. To help to motivate pronoun production, the model had a DEFPRO unit in the Concept layer which helped to signal pronouns and articles (more detail about the implementation is available in the appendix in Chang, Dell and Bock 2006 and in Chang 2009). Note that this raises the important point that argument omission and pronouns are quite different cross-linguistic means of dealing with recoverable referents and may need to have their own motivations in models of production.

The meaning pathway instantiated the message in role-concept links, but these links were not accessible to the SRN and therefore the SRN did not know how many roles were in the message. Since the number of roles is important for determining the structure of the sentence, this information was provided to the SRN through a separate Event-semantics layer (top of Figures 2 and 3). This layer had special role units for agents, patients, and goals and these units became activated depending on the number of roles in the message (e.g., one role unit for *the dog sleeps*, two role units for *the girl chased the boy*, and three role units for *I gave the book to the man*). The relative activation of these event-semantic role units was associated with different structural alternatives and information status like focus and topicalization. For example, an English passive was associated with the situation where the patient event-semantic unit was more activated than the agent event-semantic unit (e.g., *the cat was chased by the dog*). In Japanese, this same situation was associated with a scrambled sentence (e.g., *neko o inu ga oikaketa*). This implemented the assumption that there is some discourse/saliency motivation for why people use non-canonical structures like passives or scrambled structures (Bock and Irwin 1980; Ferreira and Yoshita 2003; Prat-Sala and Branigan 2000).

The event-semantic units played an important role in dealing with multi-clause utterances (see Chang 2009). For example, in Japanese utterance (6), there are two clauses with their own messages.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| (6) | *otoko* | *ga* | ***otokonoko*** | ***ga*** | ***mituketa*** | ***onnanoko*** | *ni* | *keeki* | *o* | *ageta.* |
| | man | NOM | boy | NOM | found | girl | DAT | cake | ACC | gave |

‘The man gave **the girl that the boy found** the cake.’

The main clause *otoko ga onnanoko ni keeki o ageta* means ‘the man gave the girl the cake’ and the embedded clause *otokonoko ga mituketa onnanoko* means ‘the boy found the girl’. Importantly, GIRL is a part of both messages, but she is the recipient of the GAVE action and the patient of the FOUND action. To deal with these multi-clause messages, it was assumed that there were two sets of roles in the role layer and two sets of role units in the event-semantics. Since GAVE has three arguments, the agent, patient, and goal units were activated in the main clause event-semantics units. Since FOUND has two arguments, the agent and patient units were activated in the embedded clause event-semantics units. But importantly, it was also necessary to signal that the GIRL recipient of the main clause was the same GIRL that was the patient of the embedded clause. This was achieved by activating a special event-semantic unit for the main-clause-recipient and embedded-clause-patient binding. Therefore, the number of roles in each clause and the relation of these roles to each other was encoded in the event-semantics and these units helped the SRN to select appropriate syntactic structures for conveying this information.

The final part of the Dual-path model is particularly important for incremental planning. Word choices early in sentences can influence later word choices. For example, if one is trying to convey the meaning behind *the dog chased the cat* and one starts with the phrase that describes the patient (*the cat*), then one needs to recognize that a passive is necessary to convey the appropriate message (e.g., *the cat was chased by the dog*). To do this, a reverse role-concept system was required to map from the previous word to its concept and then to its role in the particular message. Since this reverse system operated in the comprehension direction, we prefaced the units in this reverse network with *Comp* to distinguish them from their counterparts in the forward part of the network. When word *cat* was activated in the PrevWord layer, then the concept CAT was activated in the CompConcept layer and its role in the present message (e.g., patient) was activated in the CompRole layer. The CompRole roles then connected to the SRN and the model learned that sentences that start with the patient were likely to be produced in the passive. The mapping through the CompConcept-CompRole message allowed the model's syntactic choices to be sensitive to lexical items that the model has produced previously.

It is important to mention that the inclusion of a message means that this model has structural abilities that go beyond some of the limitations of many connectionist

and statistical learners (Chomsky 1975). For example, linguistic theories have posited an auxiliary inversion rule that insures that the "is" in (7a) is from the matrix verb "running". SRN models learn surface regularities and may learn from utterances like (7b) a general bias towards omitting the auxiliary between *that* and the verb, and this bias can interfere with the learning the auxiliary inversion rule. The Dual-path model on the other hand can use the message to learn the right regularities. For example, utterances like (7c) are associated with messages like RUN(BOY, PROGRESSIVE,YESNO?) and the model must learn that PROGRESSIVE+YESNO? triggers auxiliary fronting in English (but not Japanese).

(7) a. *Is the boy that is chasing the girl running?*
b. *Is the boy that chases the girl running?*
c. *Is the boy running?*

When complex messages are presented like RUN(BOY,PROGRESSIVE,YESNO?) + CHASE(BOY,GIRL,PROGRESSIVE), the model can use the message to map the auxiliary from the main clause RUN predicate, rather than the closer auxiliary from embedded clause CHASE predicate. These meaning representations come from non-linguistic information about events (the RUN and CHASE predicates can be distinguished by different actions necessary to accomplish each predicate). Therefore, the message representations in the Dual-path model provide a way to explain some of the generalizations associated with syntactic structural knowledge and movement operations in linguistic theories (see Fitz and Chang (2008) for more about how the model explains other structural phenomena like the accessibility hierarchy).

To summarize, the Dual-path model was composed of two pathways: the sequencing pathway SRN that learned syntactic regularities and the meaning pathway that encoded semantic information. The meaning pathway had role-concept links that allowed the SRN to dynamically activate appropriate concepts at different sentence positions, as well as event-semantic information about the number of arguments in the message and the relationships between arguments in different clauses. The CompConcept-CompRole reverse message allowed the SRN to recognize the role of previously produced words, which allowed the model to deal with structural alternations (e.g., the difference between passive and active constructions). All of the links in Figures 2 and 3 (except for the message links) were learned from the message-sentence training pairs. Thus, the network architecture and learning mechanism are the main innate assumptions of this model. Phonology and lexical items are assumed to be acquired by lower-level systems and the message is set by non-linguistic event understanding systems. But once these elements are available, the Dual-path model provides an input-driven learning account of the acquisition of syntactic knowledge and its interface with meaning. Although other models of incremental sentence production exist (Kempen and Hoenkamp 1987), the Dual-path model is one of the few production models that can learn to order words in

typologically-different languages like Japanese (Chang, Lieven and Tomasello 2008; Chang 2009).

# 3 Learning and adult sentence production

To examine the Dual-path model's account of psycholinguistic data, we start with a phenomenon which highlights the link between learning and processing called syntactic or structural priming (Bock 1986a; Pickering and Ferreira 2008). Structural priming is the tendency for speakers to repeat structures of previously heard sentences. For example, speakers can convey a similar meaning with a double object sentence (8a) or a prepositional dative sentence (8b). But if participants have heard another prepositional dative sentence like (8c) beforehand, they are more likely to use the same structure in their own utterances.

(8) a. *The woman gave the man a book.*
 b. *The woman gave a book to the man.*
 c. *The boy told the story to his friend.*

This phenomenon seems to occur even when words or roles are varied (Bock and Loebell 1990; Bock 1989), which suggests that it depends on abstract syntactic representations (e.g., in the prepositional dative case, something like NP VERB NP PP). Critically, structural priming seems to persist over time even when ten intervening, structurally unrelated sentences are produced between prime and target (lag 10), suggesting that this reflects a long term change in the likelihood of using a structure (Bock, Dell, Chang and Onishi 2007; Bock and Griffin 2000). Priming is normally a short-term phenomenon which disappears within a second (Levelt, Roelofs and Meyer 1999) and therefore, the long duration of structural priming is surprising and suggests that priming must be due to a type of learning mechanism. Since speakers are often not aware that they are being primed, this learning is not explicit learning like when adults are taught a language in school, but a type of implicit learning like when children acquire their first language from mere exposure to language.

To explore the idea that structural priming was a type of implicit language learning, Chang, Dell and Bock (2006) examined whether the syntax learning mechanism in the Dual-path model could also explain priming. By leaving the learning mechanism ON during the processing of the prime sentence, they found that the changes in internal representations due to the model's language learning algorithm increased the model's tendency to use the same structure to describe the target message; that is, the model demonstrated structural priming. Priming can be seen in the difference in the percentage production of a structure when preceded by the

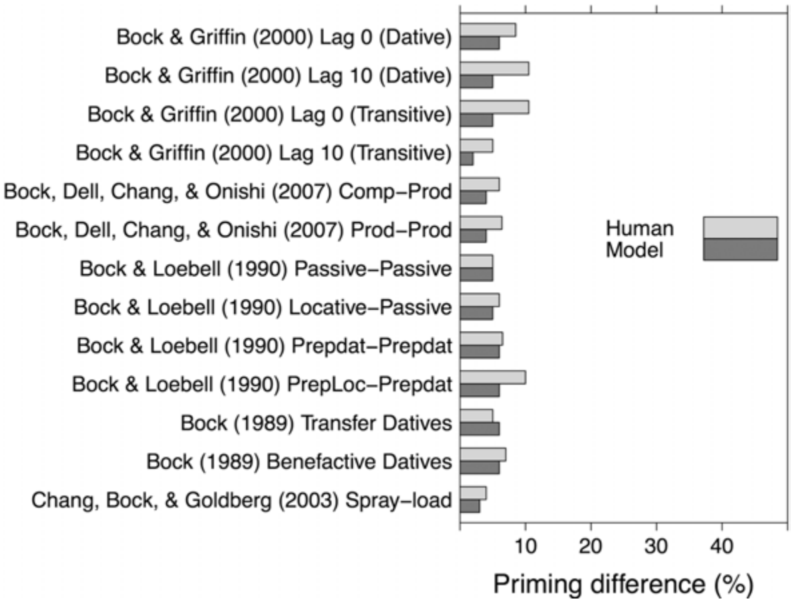



**Figure 4:** Percentage difference for human and model structural priming (Chang, Dell and Bock 2006)

same structure prime versus a different structure prime (*priming difference*). For example, 5% priming difference might arise if passive structures are produced at 25% after passive primes and 20% after active primes. Figure 4 shows the human and model priming differences for a range of studies. The model's behavior was said to be similar to human behavior if the model showed significant positive priming difference when humans also have significant positive priming difference (the exact magnitude of priming is not explained by the model). The critical test of structural priming as learning comes from the ability to model the human results in Bock and Griffin (2000, dative/transitive lag 0 and 10 in top of Figure 4), where they showed similar structural persistence when prime and target were adjacent (no intervening sentences, lag 0) and when they were separated by ten intervening sentences (lag 10).

Connectionist models learn their syntactic representations and they need to be tested to see if they approximate human representations. The similarity of the model's learned representation to those in humans can be seen in its ability to explain several priming studies that support the idea that priming involves abstract syntactic structures (Figure 4), such as insensitivity to overlap in roles (Bock and Loebell 1990), insensitivity to closed class words (Bock 1989), insensitivity to morphological overlap (Pickering and Branigan 1998), shared representations between production and comprehension (Bock et al. 2007), and sensitivity to the order of thematic roles in some constructions (Chang, Bock and Goldberg 2003). Thus the

same learning mechanism that can explain the small changes that support the persistence of priming in adults can also explain the larger changes that occur in language development where syntactic structures are learned.

The model's account of priming also predicts that, since Japanese speakers learn Japanese syntax using the same learning mechanism as English speakers, their language processing should also exhibit structural priming effects (Arai 2012). In addition, it predicts that because language users learn different syntactic representations, structural priming in English and Japanese can have different properties. One study that provides evidence for both of these claims is a study by Yamashita, Chang and Hirose (2003). They compared canonical order primes like (9a) with scrambled primes like (9b) and found that speakers were more likely to change a given order to canonical order when the prime was canonical compared to when it was scrambled. This demonstrated that priming occurs for structures that differ only in scrambling. They also included a locative condition with a transitive verb (9c).

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (9) | a. | *CEO* | *wa* | *kaigosisetu* | *ni* | *wagonsya* | *o* | *kihusita.* |
| | | | TOP | retirement home | DAT | wagon | ACC | donated |

'The CEO donated the wagon to the retirement home.'

b. *CEO wa wagonsya o kaigosisetu ni kihusita.*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| c. | *CEO* | *wa* | *kaigosisetu* | *ni* | *wagonsya* | *o* | *tometa.* |
| | | TOP | retirement home | DAT | wagon | ACC | parked |

'The CEO parked the wagon at the retirement home.'

Bock and Loebell (1990) found that English locative structures prime dative structures, but the Yamashita et al. study showed less priming in the Japanese locative condition. These results suggest that priming in English may depend more on surface structure than in Japanese, and this difference would be difficult to explain within a universal processing account. But if priming is due to learning and Japanese learners have learned different representations than English speakers, then these priming differences can be explained within the same common learning mechanism.

## 4 Heavy NP shift in English and Japanese

The model's tight link between language acquisition and language processing provides a way to understand cross-linguistic differences in language processing. One example of cross-linguistic variation comes from the phenomenon of heavy NP shift (Arnold, Losongco, Wasow and Ginstrom 2000; Hawkins 1994, 2004; Ross 1967). Heavy NP shift in English is a tendency to prefer configurations where long phrases

are placed later in sentences. For example, speakers might change sentence (10a) into (10b) (dative shift), where the long phrase *the woman that he met last week* is at the end of the sentence.

(1) a. *The man gave the woman that he met last week the book.*
b. *The man gave the book to the woman that he met last week.*

The English short-before-long bias is consistent with an incremental sentence planning architecture where short phrases are easier to plan (more accessible) and hence are produced earlier in sentences (Kempen and Hoenkamp 1987; Wasow 2002). But while English speakers have a short-before-long bias, Japanese speakers have the opposite bias to place long phrases before short phrases (long-before-short) (Hawkins 1994; Yamashita and Chang 2001). Children also show the English- and Japanese-specific direction of heavy NP shift, which suggests that it is present from fairly early in development (de Marneffe, Grimm, Kirby and Bresnan 2012; Hakuta 1981). Yamashita and Chang (2001) demonstrated the Japanese long-before-short bias in an experimental study where participants had to create a sentence from a set of phrases. They varied the length of subject and object phrases for transitive verbs and direct object (DO) and indirect object (IO) for dative verbs.

**Table 1:** Experimental conditions in experiment 2 in Yamashita and Chang (2001)

| Condition | Japanese | English |
|---|---|---|
| All-Short | *Masako wa otoko ni keeki o haitatusita.* | Masako delivered the cake to the man. |
| Long-IO | *Masako wa sinbun de syookaisareta otoko ni keeki o haitatusita.* | Masako delivered the cake to the man who was introduced in the newspaper. |
| Long-DO | *Masako wa otoko ni sinbun de syookaisareta keeki o haitatusita.* | Masako delivered the cake which was introduced in the newspaper to the man. |

Dative conditions are shown in Table 1. The canonical order for dative verbs in Japanese is SUBJECT-IO-DO (*Masako wa otoko ni keeki o haitatusita*) and speakers in this study used this order the majority of the time in all conditions. But their willingness to use the shifted order where indirect and direct objects were flipped (*Masako wa keeki o otoko ni haitatusita*) varied by condition. In the All-short condition, they produced this shifted order only 5.94% of the time. When the direct object was long, the use of shifted order increased to 37.8% as this order shifted the long phrase earlier in sentence (11a). But when the indirect object was long and the shifted order put the long phrase later in sentence (11b), they were reluctant to use this order, doing so at a rate of 2.73%.

(11) a. *Masako wa sinbun de syookaisareta keeki o otoko ni haitatusita.*
‘Masako delivered to the man the cake which was introduced in the newspaper.’

b. *Masako wa keeki o sinbun de syookaisareta otoko ni haitatusita.*
‘Masako delivered the cake to the man who was introduced in the newspaper.’

These results and the authors’ similar findings for transitives show that Japanese speakers have a long-before-short bias. Hence, Japanese and English speakers have opposite biases in their placement of long phrases and theories of sentence production need to take into account this cross-linguistic difference.

One account of this cross-linguistic difference is that speakers are trying to minimize the distance between verbs and their arguments (Hawkins 1994; 2004). This account can explain the difference in the direction of heavy NP shift, because the verb position is different in English and Japanese. For example in heavy NP shifted sentence (12a), the distance between the verb and the direct object argument (*keeki o*) in words is 2 and the distance to the indirect object argument (*otoko ni*) is 0 words (treating particles like *ni* and *o* as words).

(12) a. *Masako wa sinbun de syookaisareta keeki o otoko ni haitatusita.*

b. *Masako wa otoko ni sinbun de syookaisareta keeki o haitatusita*

But in canonical sentence (12b), the distance to the direct object is 0 words, but the indirect object is separated from the verb by 5 words. The total distance is higher for the canonical sentence compared to the shifted sentence and Hawkins would argue that this is used to select the shifted structure. But the verb-argument minimization account is challenging for incremental theories of sentence production, because it requires that speakers plan multiple whole sentences and compute the verb-argument distances for each possible utterance. Since Japanese allows scrambling of arguments, a dative sentence has six possible orders that would need to be planned. At present, there is no evidence that speakers plan all of these possible utterances and therefore it is not clear how a verb-argument minimization account would work in incremental sentence production.

## 5 A Dual-path model account of heavy NP shift in English and Japanese

Since the Dual-path model can learn different languages and use its representations in incremental sentence production, it can be used to explore these cross-linguistic

differences in heavy NP shift. Before the model's heavy NP shift behavior is discussed, it is necessary to describe the English and Japanese language input that the model used for acquiring these languages (Chang 2009). The Dual-path model used message-sentence pairs as input, and the messages for the English and Japanese sentences were equated so that both languages were equally complex in terms of the meanings being conveyed. Each language had intransitive, transitive, and dative events. English passives were assumed to have the same meaning as scrambled Japanese sentences. For instance, (13a) and (13b) had the same message with event-semantic features that signaled that the patient was more activated.

(13) a. *The cat was chased by the dog.*

b.

| | | | | |
|---|---|---|---|---|
| *neko* | *o* | *inu* | *ga* | *oikaketa.* |
| cat | ACC | dog | NOM | chased |

English double object datives had the same meaning as canonical Japanese utterances (e.g., (14a) = (14b)) and prepositional dative messages were the same as the messages for scrambled Japanese datives (e.g., (15a) = (15b)).

(14) a. *The man gave the woman the book.*

b.

| | | | | | | |
|---|---|---|---|---|---|---|
| *otoko* | *ga* | *onna* | *ni* | *hon* | *o* | *ageta.* |
| man | NOM | woman | DAT | book | ACC | gave |

(15) a. *The man gave the book to the woman.*

b.

| | | | | | | |
|---|---|---|---|---|---|---|
| *otoko* | *ga* | *hon* | *o* | *onna* | *ni* | *ageta* |
| man | NOM | book | ACC | woman | DAT | gave |

English pronouns were realized as omitted arguments in Japanese (e.g., (16a) = (16b)) and there was event-semantic information that signaled that there were three arguments, but the agent was recoverable from the discourse.

(16) a. *He gave the woman the book*

b.

| | | | | |
|---|---|---|---|---|
| *onna* | *ni* | *hon* | *o* | *ageta.* |
| woman | DAT | book | ACC | gave |

Finally, inputs for both languages contained relative clauses, which in English occur after the noun phrase that they modify and in Japanese occur before the phrase that they modify (e.g., (17a) = (17b)).

(17) a. *The boy that the man gave the book slept.*

b.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *otoko* | *ga* | *hon* | *o* | *ageta* | *otokonoko* | *ga* | *neta.* |
| man | NOM | book | ACC | gave | boy | NOM | slept |

Ten English models were created by taking the same architecture and training them on 10 different input sets made up of 40,000 message-sentence pairs. Ten Japanese models were created by using the same messages as the English model, but generating the appropriate Japanese sentences to pair with them. Thus the models were the same in terms of the frequency and distribution of concepts in the meaning, but utterances that were paired with these meanings were language-specific. When tested on a novel set of message-sentence pairs, both models were able to show similar high levels of accuracy (grammatical output: English model 93%, Japanese model 95%). Grammaticality was measured by labeling both the model's target and output sentences with syntactic categories. Output and target sequences were compared for each sentence, and if they matched, the sentence was considered grammatical. Although the message-sentence mapping in these two languages was quite different, the model was able to learn both languages equally quickly. One reason for this was that the model was not forcing each language to fit into some universal set of categories. Instead it was trying to predict language-specific utterances based on meaning; the representations that were learned were therefore those that the model found to be most effective for each language.

To examine the Dual-path model's heavy NP shift behavior, it was given dative messages in which the patient or recipient phrase was made longer by modifying it with a relative clause (Chang 2009). The model then produced the utterances and the order of short and long phrases was examined. The English version produced more recipient-before-patient orders when the patient was long than when the recipient was long (e.g., long patient (18a) versus long recipient (18b)).

(18) a. *The man gave the girl <u>the cake that the boy found.</u>*

b. *The man gave <u>the girl that the boy found</u> the cake.*

On the other hand, the Japanese model produced more recipient-before-patient orders when the recipient was long than when the patient was long. The human and model data are shown in Figure 5 with the dependent measure as the difference in the production of recipient-before-patient order depending on whether the recipient was long or the patient was long (the Japanese human data comes from the dative items in Yamashita and Chang 2001, and the English human data were created by averaging the values in Figure 8 of Arnold et al. 2000). The English difference scores in Figure 5 are negative numbers, because production of the recipient-before-patient double object was lower when the recipient was long (long-before-short) than when the patient was long (short-before-long). The Japanese results in Figure 5 are positive numbers, because the recipient-before-patient canonical order was more likely produced when the recipient was long (e.g., (19a)) than when the patient was long (e.g., (19b)).

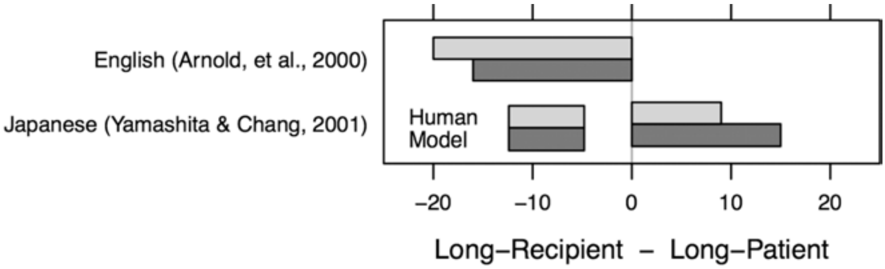



**Figure 5:** Percentage difference in use of Recipient-before-Patient order in Japanese/English human and model heavy NP shift behavior (Chang 2009)

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| (19) | a. | *otoko* | *ga* | *otokonoko* | *ga* | *mituketa* | *onnanoko* | *ni* | *keeki* | *o* | *ageta.* |
| | | man | NOM | boy | NOM | found | girl | DAT | cake | ACC | gave |
| | b. | *otoko* | *ga* | *onnanoko* | *ni* | *otokonoko* | *ga* | *mituketa* | *keeki* | *o* | *ageta.* |
| | | man | NOM | girl | DAT | boy | NOM | found | cake | ACC | gave |

These results show that the same architecture can learn a short-before-long pattern in English and a long-before-short pattern in Japanese input.

Unlike humans, computational models allow us to directly inspect the underlying representations that support their behavior. Analysis of the model's internal representations suggested that heavy NP shift behavior was due to a difference in the relative importance of meaning and surface structural information in the two languages at the *choice point* where the two word orders diverge and speakers are forced to choose between word orders. In the English dative alternation, the choice point was after the verb and at that point, the model tended to use structural cues, such as the fact that short-before-long utterances (e.g., (20a)) are more similar to simple main clause structures (e.g., (20b)) than long-before-short utterances (e.g., (20c)).

(2) a. *The man gave the girl the cake that the boy found.*

b. *The man gave the girl the cake.*

c. *The man gave the cake that the boy found to the girl.*

The high frequency of simple main clauses in the input could therefore bias the model toward the short-before-long utterance.

In contrast since Japanese is verb final and allows scrambling of all arguments of the verb, the choice point is earlier in the sentence. At these positions, it is difficult to use structural cues, since the verb occurs late in the sentence and early structural configurations are highly variable due to argument omission and scrambling. Therefore, the Japanese model preferred to use the event-semantic message information to guide its choices. Since the event-semantics signaled that a relative clause

should be produced and relative clauses go before their heads in Japanese, the model had a slight preference for producing relative clauses before main clause noun phrases, and this bias created the long-before-short bias. Thus, the model's explanation of heavy NP shift was different in each language. In English, structural similarity supported the short-before-long preference, while in Japanese, the competition between relative clauses and main clause arguments signaled by the message supported the long-before-short preference. The model demonstrated that the acquisition of a specific language can explain how the different directions of heavy NP shift arise out of different meaning and syntactic constraints at the sentence positions where structural choices were made. Unlike verb-argument minimization accounts (Hawkins 2004), the model explains heavy NP shift in different languages without generating all possible word orders.

## 6 Conceptual accessibility in English and Japanese

Heavy NP shift is unusual in that the bias in each language is in the opposite direction. More often, it is the case that processing biases in different languages are similar, but differs in ways that may be challenging to explain if the processing system is universal. One assumption of theories of language production system is that grammatical encoding involves two levels: the functional and positional levels (Bock and Levelt 1994; Garrett 1988). The functional level refers to the assignment of syntactic functions like the subject and object in a sentence. The positional level refers to the word ordering process itself. For example, in function assignment, if you are trying to convey the meaning 'the dog chased the cat', then the dog is assigned to the subject and the cat is assigned to the object function. Then in positional processing, a word order will be created which places the subject before the verb and the object after the verb. These levels were originally motivated by the distribution of speech errors in English (Garrett 1988), but they were argued to extend to sentence planning processes as well (Bock and Levelt 1994).

To understand the evidence for these levels, it is first important to understand how word accessibility influences structural planning. Accessible words are words that are easier to produce in naming studies (Bock 1982, 1986b). Many studies have found that English speakers tend to place accessible words early in sentences and this sometimes requires changes in the structure to maintain the same meaning. For example, McDonald, Bock and Kelly (1993) found that participants preferred to place words for animate concepts early in transitive sentences and this sometimes required them to use a generally less preferred structure like a passive (e.g., (21b) preferred over (21a)).

(21) a. *The sound frightened the students*
b. *The students were frightened by the sound.*
c. *The manager and the key were nowhere to be found*
d. *The key and the manager were nowhere to be found*

But when the same manipulation was done with conjunctions (e.g., (21c) and (21d)), they found that animacy did not influence word order. This difference in the influence of animacy was explained with the functional and positional levels in the production architecture. Active and passive structures differ in the element that is assigned to the subject function and this assignment takes place at the functional level in the theory. The elements in conjunctions are assigned to the same syntactic function (e.g., subject) and the ordering of these two noun phrases takes place at the positional level. Therefore, the behavioral difference between transitives and conjunctions suggested that conceptual factors like animacy could influence the functional level but not the positional level (similar results for other factors have been found, e.g., imageability; Bock and Warren 1985).

While the functional/positional distinction has been useful for explaining results in English, it is less useful for explaining behavior in Japanese. Syntactic functions in Japanese are signaled by case markers and the same case-markers are used regardless of whether a canonical or scrambled order is produced (e.g., (22a) could be said in canonical order as (22b) or scrambled order as (22c)).

(22) a. *John eats rice*

b.

| | | | | |
|---|---|---|---|---|
| *John* | *ga* | *gohan* | *o* | *taberu.* |
| John | NOM | rice | ACC | eat |

c.

| | | | | |
|---|---|---|---|---|
| *gohan* | *o* | *John* | *ga* | *taberu.* |
| rice | ACC | John | NOM | eat |

Scrambling does not change syntactic functions and therefore it should occur at the positional level (topicalization involves an additional change in the particle and this could take place at the functional level in the theory). This would predict that conceptual factors do not influence scrambling in Japanese. But in fact, it has been found that animacy and discourse status can influence Japanese scrambling (Ferreira and Yoshita 2003; Tanaka et al. 2011). To further complicate things, Tanaka et al. (2011) found that animacy does not influence the order of elements in Japanese conjunctions, which is similar to the behavior in English. Thus, production behavior for transitives and conjunctions differs in similar ways in English and Japanese, but this distinction in Japanese is difficult to explain in terms of functional and positional processing.

Can the Dual-path model, which does not have distinct functional and positional levels, explain these findings? To examine this, the English and Japanese models were given sentences to produce with animate and inanimate elements that mirrored the McDonald et al. (1993) and Tanaka et al. (2011) studies. The proportion of word order switches for each sentence type for human and model studies are shown in Figure 6. In humans and the model, English active and Japanese canonical utterances were produced accurately and word order switches were low. But English passive utterances were often switched to active utterances (word order switches were high for English Passives in Figure 6). When the given sentence had an early inanimate noun, the model was likely to switch the structure to place the animate noun earlier (English Passive Inanimate Early in Figure 6 have highest word order switches, e.g., (23a)). Similarly, the Japanese model was also likely to switch scrambled transitive utterances back to canonical utterances (Japanese Scrambled have higher word order switches than Japanese Canonical in Figure 6) and these flips were also sensitive to animacy (Japanese Scrambled Inanimate Early items have highest switch percentage, e.g., (23b)).

(23) a. *The cup was carried by the man.* → *The man carried the cup.*

b. *koppu o otoko ga motte itta.* → *otoko ga koppu o motte itta.*

The model provides a good match to the behavioral data in its sensitivity to animacy order and the way that this varies across different structures.

The model can explain why animacy influences both function assignment and scrambling, because the Prevword-CompConcept-CompRole pathway allows the model to adjust structures based on words that it has produced. For example, if the English model started a sentence with *cat* (activated in Prevword layer) and the cat was the patient in a transitive event (e.g., (24a)), then the patient role would become activated in the CompRole layer. Since the model was trained with passives where the patient role was activated early on, the SRN can use the activation of the patient role in the CompRole layer to initiate a passive (e.g., (24b)) by activating the auxiliary word *is* in the position after *cat*.

(24) a. *The dog chased the cat.*

b. *The cat is chased by the dog.*

c. *Neko o inu ga oikaketa.*
cat ACC dog NOM chased

The same CompRole signal yields different results in Japanese due to the Japanese-specific scrambling representations in the SRN. The Japanese model learned that the patient role activation in the CompRole layer at this position signals a scrambled structure (e.g., (24c)), so it will activate the *o* object particle in the Word layer at the

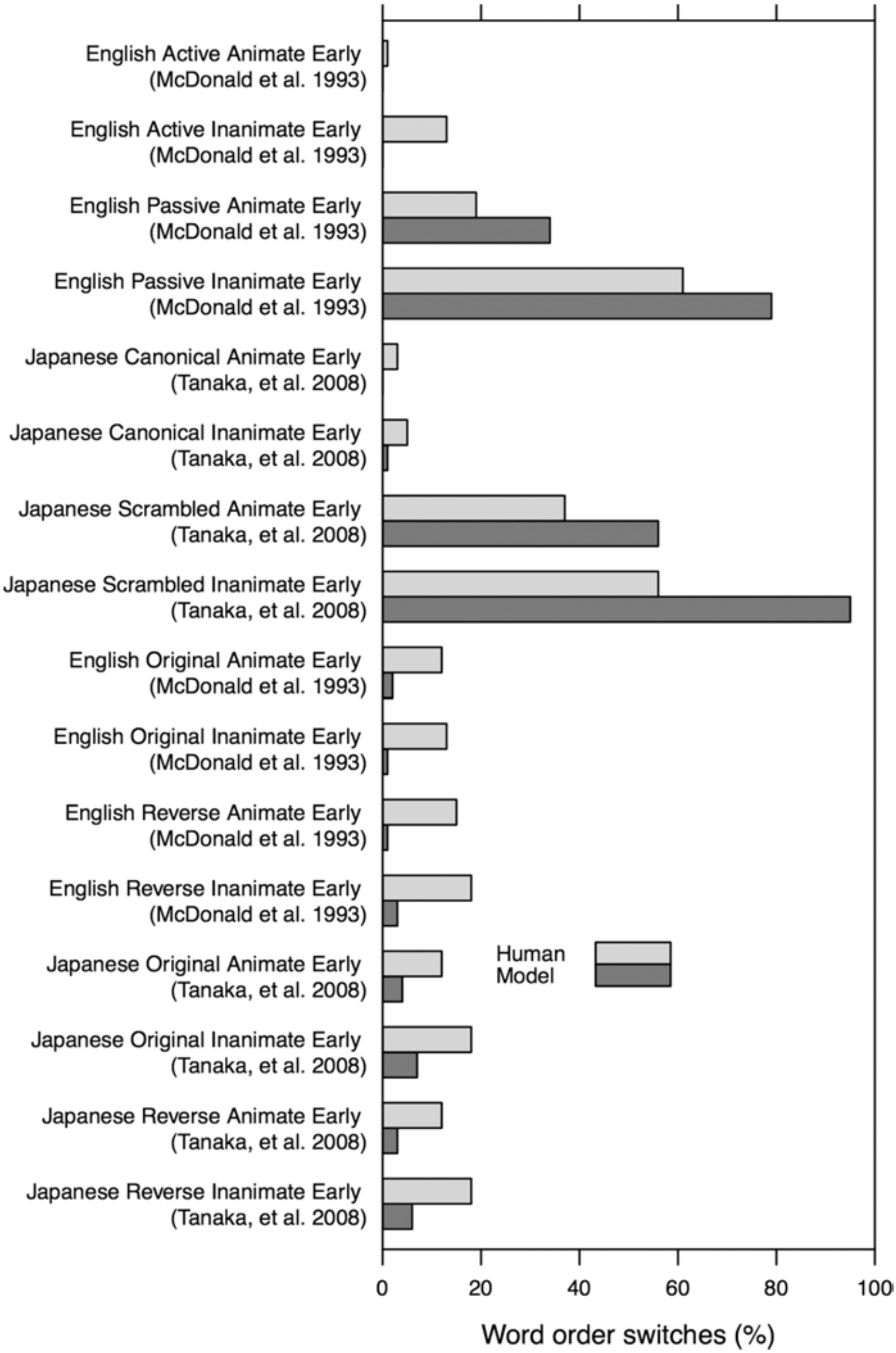



**Figure 6:** Percentage difference in word order switches for human and model accessibility behavior (Chang 2009)

position after *neko*. Thus, any factor that makes a word more likely to be produced early can influence the model's structural choices and this influence will be felt regardless of whether the language uses syntactic function assignment or scrambling to make these words prominent.

It is thought that conceptually-accessible words are more accessible because the concepts themselves are more enriched or salient. But the existence of synonyms with similar meanings but different accessibility (e.g., dog, canine) demonstrates that accessibility is partially learned. This is especially clear in the Dual-path model, where accessibility is due to the learned links between concepts and words. To create variation in accessibility, it is necessary to vary the input such that more accessible words occur earlier in sentences than less accessible words. The model's preference for placing animate elements before inanimate elements in Figure 6 (Inanimate Early items have high switch rates) is due to learning from the input distribution to expect more animate words earlier in sentences. Since accessibility is partially learned from language input, it is possible that the strength of animacy's influence on word order can vary across languages.

Another feature of accessibility that needs to be explained is the lack of an influence of animacy in conjunctions (McDonald et al. 1993; Tanaka et al. 2011). Speakers did not switch the conjunctions (25a) to the animate early order (25b). These sentences can appear strange to Japanese speakers, because verbs in Japanese are very selective in terms of animacy (Rispoli 1989) and hence mixed animacy conjunctions are awkward. Nonetheless, Tanaka et al.'s participants were able produce these sentences in the given order in his study.

| | | | | | | |
|---|---|---|---|---|---|---|
| (25) | a. | *hune* | *to* | *ryoosi* | *ga* | *ugoiteiru.* |
| | | boat | and | fisherman | NOM | moving |
| | b. | *ryoosi to fune ga ugoiteiru.* | | | | |

The model exhibited this behavior as well, as can be seen in the bottom half of Figure 6 (see English/Japanese Original/Reverse Animate/Inanimate Early). The reason that animacy had little effect on these structures was because people were unlikely to change the word order of conjunctions in both languages. Word order switches in the bottom half of Figure 6 were low compared to the passives in the top half. To explain why participants had a good memory for the order of elements in conjunctions, it was assumed that the event semantics had prominence features for the two elements in the conjunction that help to signal the given word order and that the model had to learn to associate these features with language-specific word orders. The difference between passives and conjunctions came from the asymmetry in frequency in passives, which were much less frequent than actives. This made it harder for the model to learn to use the prominence features to control passives and hence passives were more sensitive to accessibility.

The Dual-path model's account of accessibility phenomena differs from the traditional levels-based account of accessibility, which fixes certain kinds of operations to particular levels (e.g., scrambling is positional processing). Instead, the model argues that learned knowledge about words and message prominence interact to guide word order (see Prat-Sala and Branigan 2000, for evidence in support of these two mechanisms in Spanish). The model learned that animate words tended to appear before inanimate words, because this was reflected in the input. It also learned to use prominence information to guide word order and this knowledge was strong when both orders were equally likely in the input (e.g., conjunctions). When one structure was less frequent (e.g., passive, scrambled), then prominence information had less of an effect and the animacy could then influence the structure. Once the model selected earlier words, then the Prevword-CompConcept-CompRole signaled to the SRN the role of these words and the SRN used its English and Japanese knowledge to create either a passive or a scrambled sentence.

## 7 Japanese language acquisition in the Dual-path model

Although this chapter focuses on the influence of learning on adult sentence processing, there are also predictions for language acquisition. Most importantly, the Dual-path model argues that language acquisition does not need to start with innate representational knowledge such as linguistic parameters or innate categories. Rather, it argues that the acquisition of syntactic representations is easier when these representations are not required to conform to universal types. For example, if a syntactic theory has an innate adjective category, then English speakers can learn to map all modifiers to this category and place them before nouns as in (26a), (27a), and (28a). But Japanese speakers cannot use this innate adjective category, because different modifiers have different syntactic properties as shown by the capitalized elements as in (26b), (27b), and (28b).

(26) a. *the red book*
b. *aoi hon* (*aoi* is an *i*-adjective)

(27) a. *the pretty book*
b. *kirei NA hon* (*kirei* is a *na*-adjective)

(28) a. *the green book*
b. *midori NO hon* (*midori* is a noun)

It is challenging to explain these differences within a universal set of categories (Croft 2001). The Dual-path model does not force each language into a fixed set of universal categories, but instead tries to find the best way to predict sentences using its innate architecture, meaning representations, and learning algorithm.

One phenomenon that was initially argued to provide evidence for innate syntactic knowledge is syntactic bootstrapping (Naigles 1990). Syntactic bootstrapping is the idea that children can learn the meaning of novel verbs by using the syntactic structure of the utterance that the word appears in. Evidence in support of syntactic bootstrapping comes from a paradigm called preferential looking (Hirsh-Pasek and Golinkoff 1996), where two videos are simultaneously shown while an utterance is played. Infants/toddlers exhibit syntactic knowledge by looking at the video that matches the syntactic structure of the utterance. Typically these studies use videos that compare two actions that differ in causativity. The causative video shows a scene where one actor does some action to another actor (e.g., a girl pushes a boy into a sitting position). The non-causative video shows a scene where two actors do the same action together (e.g., a girl and a boy waving arms in a synchronous manner). The causative videos are best described by transitive utterances (e.g., (29a)) and the non-causative videos are best described by intransitive utterances (e.g., (29b)).

(29) a. *The girl is pushing the boy.*
 b. *The girl and the boy are waving.*

18- and 21-month-old English infants have been shown to prefer the matching video with transitive utterances with novel verbs suggesting that at this age they have structural representations that allow them to understand the meaning of a novel verb (Naigles 1990), and indicating that syntax was supported by innate knowledge. However, this claim has been controversial (Fisher 2002a; Naigles 2002; Tomasello 2003), because at later ages, there seems to be developmental change in the ability to produce transitive sentences in elicited production. These controversies about innate syntax depend in part on the task methodologies and one advantage of computational modeling is the way that it can make explicit how the task influences the way that representations are used.

Chang, Dell and Bock (2006) used the Dual-path model to account for the difference between elicited production and preferential looking in development. They modeled elicited production as the production of transitive utterances given a message where the whole sentence was produced correctly. Preferential looking, on the other hand, was modeled as the match between the actual next word and the word predicted by the model based on the message information in each video. This way of implementing preferential looking created a graded word-by-word match score that represented how well the sentence matched the causative or non-causative

messages. Since this measure of preferential looking was able to measure partial knowledge of the match between utterances and scenes, it was able to explain why preferential looking abilities appear before elicited production in language development. Thus the assumption that language knowledge that is learned in language acquisition must be deployable in an incremental fashion can help to explain why measures that are sensitive to partial utterances should show earlier evidence of syntactic knowledge than measures that are based on whole utterances.

The Dual-path model also made a set of novel predictions. The model found that transitive-causative mappings were easier to acquire than intransitive-non-causative mappings. This is surprising, because if syntactic knowledge is innate or easily learned, then both of these mappings should be equally available and should be evident in a forced-choice task like preferential looking. The model exhibited this earlier transitive knowledge because initially it had a bias to treat the first noun as the agent (first-noun-as-agent bias). This is because transitive agents are common subjects in the model's input and the model learned this regularity early on (Chang et al. 2006). Later experimental work confirmed these prediction in preferential looking studies (Gertner and Fisher 2012; Noble, Rowland and Pine 2011). Over development, the model lost this first-noun-as-agent bias as it learned to use non-causative messages to predict the first noun correctly. Eventually the model also shows the non-causative-intransitive mapping by learning to use the post-verbal information in the utterance to correctly map to the appropriate scene (e.g., if the verb is at the end of the sentence, then verb is non-causative). Support for the post-verbal structural bias comes from studies that show that two-year-olds can use structural cues in isolation to learn about verbs (Yuan, Fisher and Snedeker 2012). Thus an important claim arising from the model's performance is that there may be different components to preferential looking abilities, some which reflect top-down scene-based expectations and some which are driven by bottom-up distributional learning from the sentence.

If syntactic bootstrapping is a universal tool for learning verb meaning, then it should appear in different languages at the same time and should be relatively insensitive to language-specific properties. One universalist way that this ability has been construed is that children use the number of noun phrases to derive the meaning of the verb (e.g., two noun phrases are associated with a causative meaning, Fisher 2002b; Gertner and Fisher 2012) and it should be possible for children to recognize noun phrases equally well in different languages. In contrast to this universalist account, the account provided by the Dual-path model suggests that learned representations play a big part in preferential looking behavior. The first-noun-as-agent bias and the post-verbal structural bias are both learned representations that are useful for the task of next word prediction. Therefore, the model predicts greater variability in preferential looking behavior in different languages compared to an account of syntactic bootstrapping based on universal acquisition biases. In particular, languages like Japanese where all arguments of a verb can be

omitted provide more variable input for learning the representations that support preferential looking (Matsuo, Kita, Shinya, Wood and Naigles 2012; Rispoli 1989). Hence, this model would predict that preferential looking behavior in Japanese should be weaker and appear later than in English.

Although work in preferential looking in Japanese is only beginning, there are some findings that support this late development. Matsuo, et al. (2012) did a study in Japanese with stimuli and materials that were similar to Naigles (1990). They tested syntactic bootstrapping using the causative and non-causative videos and found robust transitive-causative looking behavior in 24-month olds when the utterance had case-marked particles, but not when the test utterance did not have these particles. The failure to find syntactic bootstrapping without case markers is critical, because case-markers are Japanese-specific and therefore cannot be a part of a universal syntactic bootstrapping mechanism. In contrast, Naigles (1990) found that by 21 months English toddlers could use the word order alone to map to a causative meaning. Thus, it would seem that English-learning toddlers can use structural cues like the number of noun phrases to recognize the meaning of causative verbs, but this same ability is not present in Japanese toddlers at a slightly older age. If this delay is maintained in other Japanese preferential looking studies, then that would suggest that syntactic bootstrapping may be partially learned and that would support learning-based accounts of preferential looking (Chang, Dell and Bock 2006).

As in some English studies (Gertner and Fisher 2012; Hirsh-Pasek and Golinkoff 1996), Matsuo et al. (2012) also found that children exhibited transitive-causative mappings, but not intransitive-non-causative mappings. The asymmetry between the transitive and intransitive mappings is actually predicted by the Dual-path model's account of preferential looking, where the greater frequency of transitive-causative mappings due to transitive and dative utterances in the input makes their subcomponents easier to acquire (Chang, Dell and Bock 2006). Matsuo et al. (2012) also provided corpus-based support for this construction frequency account of the transitive-causative asymmetry. They found that intransitives and transitives were about equally frequent in Japanese child-directed speech. If dative utterances are included to support the transitive-causative mapping, since they mark agents and patients like transitives, then the constructional frequency of transitives should be higher than intransitives and that can explain the behavioral asymmetry that Matsuo et al. (2012) found in Japanese.

The Dual-path model provides a learning-based account of preferential looking results that can be distinguished from a universalist syntactic bootstrapping account in its predictions for behavior in Japanese and English language acquisition. The model predicts slower development of preferential looking in languages like Japanese. It also predicts variation between transitive-causative and intransitives-non-causative mappings in response to the input, due to the fact that its preferential looking behavior is made up of multiple components such as the first-noun-as-agent bias

and the post-verbal structural bias. Further work is needed to understand how these accounts work in different languages.

## 8 Conclusions

Psycholinguistic theories of language processing have been developed from experimental data from English or similar European languages. The behavioral phenomena in these languages have been argued to reflect psychological processing mechanisms or the architecture of the language system. If these mechanisms are at work in all languages, then behavior in English and Japanese should be similar. In this chapter, various cases have been presented which show that English and Japanese speakers have different processing behaviors. This seems to militate for a theory that can use language-specific means to do language processing.

One way to formulate an account of language processing which can use different representations in different languages is to formulate a language acquisition theory that can create a system that does adult sentence processing. The Dual-path model is an example of this approach. The model's architecture and learning algorithm were shared across languages, but the resulting adult language production system differed substantially in each language. The model did not have a special heavy NP shift mechanism, but the different direction of the shift in English and Japanese models arose from differences in the model's dependence on message and structural cues at the different positions where structural choices were made in each language. Likewise, the model did not have a built-in functional and positional distinction, but rather it learned to depend on the message in different ways for transitive and conjunctions, and this gave the model behavioral distinctions that mirrored those in the functional/positional account. But since these representations were learned in each language, it was able to use different means (e.g., passivization, scrambling) to deal with variations in accessibility. Although this model is the only explicit model of Japanese sentence processing that can explain this range of behaviors in adult and child language processing using the same mechanisms used for explaining English sentence processing, there is still further work to be done. The focus of the model has been on production and acquisition data, but it is also necessary to model comprehension data, particularly eye-tracking in the visual world (Kamide et al. 2003). Also, lesioning the weights in the English model, which simulates brain damage, produces aphasia-like language patterns (Chang 2002), but it is not clear how an aphasic Japanese model would behave with case-marking and nominal verbs (e.g., 掃除をする).

The ultimate test of a model is whether it makes interesting predictions that are borne out in experimental studies. One general class of predictions that the model makes is that there will be cross-linguistic differences in language processing. For

example, if English structural priming studies are replicated in Japanese, some proportion of these studies should not yield the same results in Japanese, because the structural similarity that governs structural priming should be different in these two languages. Another set of predictions is due to the claim that language processing is shaped by language learning. If a language has more variability, such as the variability created by scrambling and argument omission, then there should be delayed or weaker performance related to that variation. This is the basis for the prediction that Japanese children should show some delay within preferential looking studies relative to English children, because the variability in the input creates weaker incremental representations. By forging a tighter link between acquisition and processing in different languages by using explicit models[1], we can have a richer and deeper understanding of the nature of cross-linguistic language use.

## Acknowledgments

Part of this research was supported by NTT Communication Science Laboratories, NTT Corp, Kyoto Japan. Thanks to Marc Delcroix, Hartmut Fitz, Cynthia Fisher, Marius Janciauskas, Tessei Kobayashi, Tadahisa Kondo, Mineharu Nakayama, Katherine Twomey, and Hiroko Yamashita for helpful comments and discussions.

## References

Altmann, Gerry T. M. and Yuki Kamide. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73(3). 247–264.

Ans, Bernard, Serge Carbonnel and Sylviane Valdois. 1998. A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review* 105(4). 678–723.

Arai, Manabu. 2012. What can head-final languages tell us about syntactic priming (and vice versa)? *Language and Linguistics Compass*, *6*(9), 545–559. doi:10.1002/lnc3.353

Arnold, Jennifer E., Anthony Losongco, Thomas Wasow and Ryan Ginstrom. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76(1). 28–55.

Baxt, William G. and Joyce Skora. 1996. Prospective validation of artificial neural network trained to identify acute myocardial infarction. *The Lancet* 347(8993). 12–15. doi:10.1016/S0140-6736(96)91555-X.

Bock, Kathryn. 1982. Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review* 89(1). 1–47.

Bock, Kathryn. 1986a. Syntactic persistence in language production. *Cognitive Psychology* 18(3). 355–387.

**1** More information and code for the models can be download from http://sites.google.com/site/sentenceproductionmodel/

Bock, Kathryn. 1986b. Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12(4). 575–586.
Bock, Kathryn. 1989. Closed-class immanence in sentence production. *Cognition* 31(2). 163–186.
Bock, Kathryn, Gary S. Dell, Franklin Chang and Kristine H. Onishi. 2007. Persistent structural priming from language comprehension to language production. *Cognition* 104(3). 437–458.
Bock, Kathryn and Zenzi M. Griffin. 2000. The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General* 129(2). 177–192.
Bock, Kathryn and David E. Irwin. 1980. Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior* 19(4). 467–484.
Bock, Kathryn and Willem J. M. Levelt. 1994. Language production: Grammatical encoding. In Morton Ann Gernsbacher (ed.), *Handbook of Psycholinguistics*, 945–984. Orlando, FL: Academic Press.
Bock, Kathryn and Helga Loebell. 1990. Framing sentences. *Cognition* 35(1). 1–39.
Bock, Kathryn and Richard K. Warren. 1985. Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21(1). 47–67.
Chang, Franklin. 2002. Symbolically speaking: A connectionist model of sentence production. *Cognitive Science* 26(5). 609–651.
Chang, Franklin. 2009. Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language* 61(3). 374–397.
Chang, Franklin, Kathryn Bock and Adele E. Goldberg. 2003. Can thematic roles leave traces of their places? *Cognition* 90(1). 29–49.
Chang, Franklin, Gary S. Dell and Kathryn Bock. 2006. Becoming syntactic. *Psychological Review* 113 (2). 234–272.
Chang, Franklin, Elena Lieven and Michael Tomasello. 2008. Automatic evaluation of syntactic learners in typologically-different languages. *Cognitive Systems Research* 9(3). 198–213.
Choi, Youngon and John C. Trueswell. 2010. Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology* 106(1). 41–61. doi:10.1016/j.jecp.2010.01.003.
Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
Chomsky, Noam. 1975. Reflections on language. *New York: Pantheon* 212.
Chomsky, Noam and Howard Lasnik. 1993. The theory of principles and parameters. *Syntax: An international handbook of contemporary research* 1. 506–569.
Christiansen, Morten H. and Nick Chater. 2001. Finite models of infinite language: A connectionist approach to recursion. *Connectionist psycholinguistics*, 138–176. Westport, CT: Ablex Publishing.
Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford, UK: Oxford University Press.
Culicover, Peter W. 1997. *Principles and parameters: An introduction to syntactic theory*. Oxford University Press, USA.
De Marneffe, Marie-Catherine, Scott Grimm, Inbal Arnon, Susannah Kirby and Joan Bresnan. 2012. A statistical model of grammatical choices in children's production of dative sentences. *Language and Cognitive Processes*, *27*(1), 25–61.
Dell, Gary S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93(3). 283–321.
Dell, Gary S., Franklin Chang and Zenzi M Griffin. 1999. Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, *23*(4), 517–542.
Dell, Gary S., Cornell Juliano and Anita Govindjee. 1993. Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, *17*(2), 149–195.
Dell, Gary S., Myrna F. Schwartz, Nadine Martin, Eleanor M. Saffran and Deborah A. Gagnon. 1997. Lexical access in aphasic and nonaphasic speakers. *Psychological Review* 104(4). 801–838.

Dominey, Peter Ford, Michel Hoen and Toshio Inui. 2006. A neurolinguistic model of grammatical construction processing. *Journal of Cognitive Neuroscience*, *18*(12), 2088–2107.

Elman, Jeffrey L. 1990. Finding structure in time. *Cognitive Science* 14(2). 179–211.

Elman, Jeffrey L. 1993. Learning and development in neural networks: The importance of starting small. *Cognition* 48(1). 71–99.

Enke, David and Suraphan Thawornwong. 2005. The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications* 29(4). 927–940. doi:10.1016/j.eswa.2005.06.024.

Evans, Nicholas and Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32(5). 429–492.

Ferreira, Victor S. and Hiromi Yoshita. 2003. Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research* 32(6). 669–692.

Fisher, Cynthia. 2002a. The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition* 82(3). 259–278.

Fisher, Cynthia. 2002b. Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science* 5(1). 55–64.

Fitz, Hartmut and Franklin Chang. 2008. The role of the input in a connectionist account of the accessibility hierarchy in development. In Harvey Chan, Heather Jacob and Enkeleida Kapia (eds.), *Proceedings of the Boston University conference on language development*, vol. 32, 120–131. Somerville, MA: Cascadilla Press.

Fodor, Jerry A. and Zenon W. Pylyshyn. 1988. Connectionism and cognitive architecture: A critical analysis. *Cognition* 28(1–2). 3–71.

Garrett, Merrill F. 1988. Processes in language production. In Frederick J. Newmeyer (ed.), *Linguistics: The Cambridge survey, III. Language: Psychological and biological aspects*, 69–96. Cambridge, UK: Cambridge University Press.

Gertner, Yael and Cynthia Fisher. 2012. Predicted errors in children's early sentence comprehension. *Cognition* 124(1). 85–94. doi:10.1016/j.cognition.2012.03.010.

Hakuta, Kenji. 1981. Grammatical description versus configurational arrangement in language acquisition: The case of relative clauses in Japanese. *Cognition* 9. 197–236.

Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge, UK: Cambridge University Press.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. New York: Oxford University Press.

Hirsh-Pasek, Kathy and Roberta M. Golinkoff. 1996. *The origins of grammar*. Cambridge, MA: The MIT Press.

Ijuin, Mutsuo, Takao Fushimi, Karalyn Patterson, Naoko Sakuma, Masayuki Tanaka, Itaru F. Tatsumi, Tadahisa Kondo and Shigeaki Amano. 2000. A connectionist approach to naming disorders of Japanese in dyslexic patients. *Proceedings of International Conference on Spoken Language Processing 2000*, Vol. 2, 32–37.

Jackendoff, Ray. 1992. *Semantic structures*. Cambridge, MA: The MIT Press.

Joe, Kazuki, Yoshihiro Mori, and Sei Miyake. 1990. Construction of a large-scale neural network: Simulation of handwritten Japanese character recognition on NCUBE. *Concurrency: Practice and Experience 2*(2), 79–107. doi:10.1002/cpe.4330020202

Kamide, Yuki, Gerry T. M. Altmann and Sarah L. Haywood. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language* 49(1). 133–156. doi:10.1016/S0749-596X(03)00023-8.

Kempen, Gerard and Edward Hoenkamp. 1987. An incremental procedural grammar for sentence formulation. *Cognitive Science* 11(2). 201–258.

Levelt, Willem J. M., Ardi Roelofs and Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22(01). 1–38.

Matsuo, Ayumi, Sotaro Kita, Yuri Shinya, Gary C. Wood and Letitia R. Naigles. 2012. Japanese two-year-olds use morphosyntax to learn novel verb meanings. *Journal of Child Language* 39(3). 637–663.

Mazuka, Reiko. 1998. *The development of language processing strategies: A cross-linguistic study between Japanese and English*. Lawrence Erlbaum.

McDonald, Janet L., Kathryn Bock and Michael H. Kelly. 1993. Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology* 25(2). 188–230.

Miyamoto, Tadao. 1999. *The light verb construction in Japanese: The role of the verbal noun*. Amsterdam: John Benjamins Publishing.

Mori, Yoshihiro and Kazuhiko Yokosawa. 1989. Neural networks that learn to discriminate similar Kanji characters. In David S. Touretzky (ed.), *Advances in neural information processing systems 1*, 332–339. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Naigles, Letitia R. 1990. Children use syntax to learn verb meanings. *Journal of Child Language* 17 (02). 357–374.

Naigles, Letitia R. 2002. Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition* 86(2). 157–199.

Negishi, Michiro. 2006. Connectionist models. In Mineharu Nakayama, Reiko Mazuka and Yasuhiro Shirai (eds.), *The handbook of East Asian psycholinguistics*, *Vol. 2, Japanese*, 315–322. Cambridge, UK: Cambridge University Press.

Noble, Claire H., Caroline F. Rowland and Julian M. Pine. 2011. Comprehension of argument structure and semantic roles: Evidence from English-learning children and the forced-choice pointing paradigm. *Cognitive Science* 35(5). 963–982. doi:10.1111/j.1551-6709.2011.01175.x.

Oppenheim, Gary M., Gary S. Dell and Myrna F. Schwartz. 2010. The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition* 114(2). 227–252.

Palmer-Brown, Dominic, Jonathan A. Tepper and Heather M. Powell. 2002. Connectionist natural language parsing. *Trends in Cognitive Sciences* 6(10). 437–442.

Pickering, Martin J. and Holly P. Branigan. 1998. The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language* 39(4). 633–651.

Pickering, Martin J. and Victor S. Ferreira. 2008. Structural priming: A critical review. *Psychological Bulletin* 134(3). 427–459.

Pinker, Steven. 1984. *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Plaut, David C. and Christopher T. Kello. 1999. The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In Brian MacWhinney (ed.), *The emergence of language*, 381–415. Mahwah, NJ: Lawrence Erlbaum Associates.

Plaut, David C., James L. McClelland, Mark S. Seidenberg and Karalyn Patterson. 1996. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review* 103(1). 56–115.

Plunkett, Kim and Patrick Juola. 1999. A connectionist model of English past tense and plural morphology. *Cognitive Science* 23(4). 463–490. doi:10.1207/s15516709cog2304_4.

Pomerleau, Dean A. 1993. Knowledge-based training of artificial neural networks for autonomous robot driving. In Jonathan H. Connell and Sridhar Mahadevan (eds.), *Robot learning*, 19–43. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Prat-Sala, Mercè and Holly P. Branigan. 2000. Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish,. *Journal of Memory and Language* 42(2). 168–182.

Rispoli, Matthew. 1989. Encounters with Japanese verbs: caregiver sentences and the categorization of transitive and intransitive action verbs. *First Language* 9(25). 57–80. doi:10.1177/014272378900902506.

Rogers, Timothy T. and James L. McClelland. 2004. *Semantic cognition: A parallel distributed processing approach*. MIT Press.

Ross, John Robert. 1967. Constraints on variables in syntax. Cambridge, MA: MIT Dissertation.

Snedeker, Jesse and John C. Trueswell. 2004. The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology* 49(3). 238–299.

Tanaka, Mikihiro N., Holly P. Branigan, Janet F. McLean and Martin J. Pickering. 2011. Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language* 65(3). 318–330.

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tsuzuki, Takashi. 1996. A connectionist model of lexical ambiguity resolution in sentence comprehension. *Japanese Psychological Review 39*(3), 273–294.

Van Gompel, Roger P. G. and Martin J. Pickering. 2007. Syntactic parsing. In Gareth Gaskell (ed.), *The Oxford handbook of psycholinguistics*, 289–307. Oxford, UK: Oxford University Press.

Vosse, Theo and Gerard Kempen. 2000. Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition* 75(2). 105–143.

Waibel, Alexander, Toshiyuki Hanazawa, Geoffrey Hinton, Kiyohiro Shikano and Kevin J. Lang. 1988. Phoneme recognition: Neural networks vs. hidden Markov models (pp. 107–110). Presented at the IEEE International Conference on Acoustic, Speech, and Signal Processing. doi:10.1109/ICASSP.1988.196523

Wasow, Thomas. 2002. *Postverbal behavior*. Stanford, CA: Center for the Study of Language and Information.

Yamashita, Hiroko and Franklin Chang. 2001. “Long before short” preference in the production of a head-final language. *Cognition* 81(2). B45–B55.

Yamashita, Hiroko, Franklin Chang and Yuki Hirose. 2003. Language-dependent aspects of structural priming. Paper presented at the Architectures and Mechanisms of Language Processing, Glasgow, UK.

Yang, Charles D. 2003. *Knowledge and learning in natural language*. New York: Oxford University Press.

Yu, Chen and Linda B. Smith. 2007. Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420.

Yuan, Sylvia, Cynthia Fisher and Jesse Snedeker. 2012. Counting the nouns: Simple structural cues to verb meaning. *Child Development* 83(4). 1382–1399. doi:10.1111/j.1467-8624.2012.01783.x.

Masatoshi Koizumi

# 13 Experimental syntax: Word order in sentence processing

## 1 Introduction

The objective of this chapter is to illustrate the types of experimental studies currently underway in the field of syntax. There are basically two types of research into sentence processing ("experimental syntax" in a broad sense): (I) research evaluating processing theories/hypotheses by testing their predictions and (II) research evaluating linguistic theories/hypotheses by testing their predictions ("experimental syntax" in a narrow sense). In Sections 2 and 3, we review several studies in categories (I) and (II), focusing on word order processing. Section 4 concludes the chapter.

## 2 From grammar to processing

In this section, we review several studies in category (I). The primary purpose of these studies is to investigate the nature of the human parser by assessing which grammatical (and non-grammatical) factors affect human sentence processing, to what extent, and in what manner. These studies, either implicitly or explicitly assuming certain grammatical theories, evaluate existing processing theories/hypotheses, and, in some cases, propose a modification or alternative.

### 2.1 Order of arguments

As in many languages, including Basque, Finnish, German, Hindi, Korean, Persian, Russian, Serbo-Croatian, Sinhalese, and Turkish, word order in Japanese is relatively free. For example, a Japanese active transitive sentence may have either of the two orders presented in (1).

| | | | | | | |
|---|---|---|---|---|---|---|
| (1) | a. | *Tomoko* | *ga* | *Taroo* | *o* | *home-ta.* |
| | | Tomoko | NOM | Taro | ACC | praise-PST |
| | | 'Tomoko praised Taro.' | | | | |
| | b. | *Taroo* | *o* | *Tomoko* | *ga* | *home-ta.* |
| | | Taro | ACC | Tomoko | NOM | praise-PST |

Similarly, the three nominal arguments in a ditransitive sentence may assume any of the six logically possible word orders. However, not all word orders are comparable. The word order in (1a), for instance, is used more frequently than that in

(1b) (Imamura and Koizumi 2011). It has also been shown, in many behavioral as well as functional brain imaging studies, that the order in (1a) is easier to process than that in (1b) (Chujo 1983; Kim et al. 2009; Kinno et al. 2008; Mazuka, Ito, and Kondo 2002; Miyamoto and Takahashi 2002; see also Hagiwara et al. 2007; Inubushi et al. 2012; Koizumi et al. 2012; Nakano, Felser and Clahsen 2002; Ueno and Kluender 2003). Thus, the word order in (1b) is more marked than the order in (1a), which is basic or unmarked. A question then arises as to what makes (1b) more marked than (1a).

From a linguistic viewpoint, there are at least three possible ways to describe the two word orders in (1) (cf. Shibatani 1977). First, they can be characterized in terms of grammatical functions such as the subject and object. In (1a), the first noun phrase (NP) *Tomoko ga* is the subject, and the second NP *Taroo o* is the object, whereas in (1b), the first NP is the object and the second NP is the subject. In Japanese generative linguistics, it is commonly assumed that subject-object-verb (SOV) word order, as in (1a), represents a basic syntactic structure, as in (2a), whereas object-subject-verb (OSV) word order, as in (1b), is a reflection of a more complex syntactic representation, as in (2b), in which the object is dislocated or moved to the left of its canonical position (Hoji 1985; Miyagawa 1989; Saito 1985; among others; see Nemoto 1999 for an overview). This type of dislocation is referred to as "scrambling." The moved object is associated with a phonetically empty element in its canonical position (represented as *t* in (2) and hereafter). The fronted constituent is often referred to as an antecedent or filler, and the empty category in the original position is called a trace or gap.

(2) Schematic mental representations of (1a, b)
a. $[_{S}$ NP-ga $[_{VP}$ NP-o V]]
b. $[_{S}$ NP-o$_1$ $[_{S}$ NP-ga $[_{VP}$ $t_1$ V]]]

If we follow the standard methodology in cognitive neuroscience, that is, assuming "all other things being equal, the more complex a representation ... [is], the longer it should take for a subject to perform any task involving the representation and the more activity should be observed in the subject's brain areas associated with creating or accessing the representation and with performing the task" (Marantz 2005: 439; see also Gibson 2000; Hawkins 2004; Pritchett and Whitman 1995; among many others), we would expect that (1a) is easier to process than (1b) because the former has the canonical SOV word order associated with a simpler syntactic structure.[1]

A second possible method used to describe the two word orders in (1) involves thematic roles such as those of the Agent and Patient. (1a) has an agent-patient order, with *Tomoko* as the agent and *Taroo* as the patient. (1b), on the other hand,

**1** Of course, other things are not always equal. In some cases, the word order of (1a) is more difficult to process than that of (1b). See Inui et al. (1998) and Nakayama and Lewis (2001), among many others.

assumes a patient-agent order. Primus (1999) argued that the agent-patient order is preferred to the patient-agent order because the patient's involvement in an event depends on the agent (and his or her actions), rather than vice versa (see also Keenan and Comrie 1977). If so, it is likely that (1a) is easier to process than (1b) because the former represents the preferred thematic role order.

A third method of characterizing the word orders in (1) involves case marking. (1a) has a nominative-accusative order, whereas (1b) has an accusative-nominative order. Just as there are dependency relations among thematic roles, there are dependency relations among cases, as well. In particular, the presence of cases other than the nominative case in a finite clause presupposes the presence of the nominative case in that clause (Shibatani 1978; Marantz 1992). Therefore, it is possible that (1a) is easier to process than (1b) because of the difference between their respective case orders.

In sum, there are three linguistically significant methods to characterize the two word orders in (1), each of which provides us with a means with which to construct a hypothesis as to why (1a) is easier to process than (1b). The three hypotheses are summarized in (3).

(3) a. The Grammatical Functions Hypothesis:
The human cognitive system is developed in such a manner that the Subject-NonSubject (S-NS) order is easier to process than the NonSubject-Subject (NS-S) order.

b. The Thematic Roles Hypothesis:
The human cognitive system is developed in such a manner that the Agent-NonAgent (A-NA) order is easier to process than the NonAgent-Agent (NA-A) order.

c. The Case Marking Hypothesis:[2]
The human cognitive system is developed in such a manner that the Nominative-NonNominative (N-NN) order is easier to process than the NonNominative-Nominative (NN-N) order.

---

**2** In ergative languages, the presence of cases other than the absolutive case presupposes the presence of the absolutive case. Therefore, the counterpart of (3c) in ergative languages should be (i).

(i) The Case Marking Hypothesis in Ergative Languages:
The human cognitive system is developed in such a manner that the Absolutive-NonAbsolutive (A-NA) order is easier to process than the NonAbsolutive-Absolutive (NA-A) order.

This hypothesis is consistent with the results of the Kaqchikel sentence processing experiment discussed in Section 3.3. It has been reported, however, that in Basque, an SOV ergative language with pro-drop, SOV (= Ergative-Absolutive-V) sentences are easier to process than corresponding OSV (= Absolutive-Ergative-V) sentences (Erdocia et al. 2009). This, together with the results of the other experiments reviewed in this chapter, suggests that in ergative languages as well as in nominative languages, the most preferred word order is the syntactically basic word order.

The three competing hypotheses in (3) all make the same prediction: that the word order in (1a) is easier to process than that in (1b). In order to evaluate their general validity, a series of behavioral and brain imaging studies has been conducted on the processing of a wider range of constructions (e.g., Kim et al. 2009; Kimura, Kim and Koizumi 2005; Kinno et al. 2008; Koizumi and Tamaoka 2004, 2006, 2010; Tamaoka et al. 2005). We now focus on one of these experimental studies, Tamaoka et al. (2005).

To investigate the primary factor that determines the cognitive load in processing alternative word orders, Tamaoka et al. (2005) conducted five reading experiments involving a sentence plausibility judgment task (Chujo 1983; Stromswold et al. 1996). In this chapter, we consider the three that are most pertinent to the hypotheses in (3). In the first experiment, active transitive sentences such as (1a, b), repeated here as (4a, b), were visually presented to the participants in random order at the center of a computer screen, where each sentence appeared as a whole at once in one line. The participants were instructed to respond as quickly and accurately as possible in deciding whether or not the sentences made sense, and they registered their responses by pressing a "yes" or "no" button. The duration between the appearance of a sentence on the screen and the button press was recorded as the reaction time. To determine whether a sentence made sense, the participant had to determine its syntactic structure as well as retrieve lexical information.

(4) a. *Tomoko ga Taroo o home-ta.*
Tomoko NOM Taro ACC praise-PST
'Tomoko praised Taro.'

b. *Taroo* $o_i$ *Tomoko ga* $t_i$ *home-ta.*
Taro ACC Tomoko NOM praise-PST

As we have seen above, all three competing hypotheses in (3) predict that (4a) is easier, and hence processed faster, than (4b). The reaction times were indeed reliably smaller for (4a) (mean reaction time = 1209 milliseconds) than for (4b) (1432 ms), consistent with the prediction as well as the results of previous studies.

The second experiment used passive sentences such as those in (5).

(5) a. *Taroo ga Tomoko ni home-rare-ta.*
Taro NOM Tomoko by praise-PASS-PST
'Taro was praised by Tomoko.'

b. *Tomoko* $ni_i$ *Taroo ga* $t_i$ *home-rare-ta.*
Tomoko by Taro NOM praise-PASS-PST

The word order of (5a) is the syntactically canonical S-NS order. The subject *Taroo* bears the thematic role of patient and is marked with the nominative case

marker *-ga*. Thus, according to the Grammatical Functions Hypothesis and Case Marking Hypothesis, (5a) should be processed faster than (5b), whereas the Thematic Roles Hypothesis predicts the opposite outcome. The results of the second experiment ((5a) = 1521 ms, (5b) = 1722 ms) supported the prediction of the Grammatical Functions Hypothesis and Case Marking Hypothesis against that of the Thematic Roles Hypothesis.

The third experiment exploited potential sentences such as those in (6).

(6) a. *Kenzi ni tyuugokugo ga yom-e-ru daroo-ka.*
Kenzi DAT Chinese NOM read-can-PRS wonder-Q
'I'm wondering if Kenzi can read Chinese.'

b. *Tyuugokugo* $ga_i$ *Kenzi ni* $t_i$ *yom-e-ru daroo-ka.*
Chinese NOM Kenzi DAT read-can-PRS wonder-Q

(6a) follows the syntactically canonical S-NS order, in which the subject is the Agent and is case-marked as dative. The Grammatical Functions Hypothesis and Thematic Roles Hypothesis predict that (6a) should be read faster than (6b). The Case Marking Hypothesis, on the other hand, foresees longer reaction times for (6a) than for (6b). The third experiment revealed that (6a) (1326 ms) was processed faster than (6b) (1542 ms), which is consistent with the Grammatical Functions Hypothesis and Thematic Roles Hypothesis, but not with the Case Marking Hypothesis.

In summary, we have seen that only the Grammatical Functions Hypothesis (3a) is consistent with the results of all three experiments described above, suggesting that the linear ordering of grammatical functions is the primary determinant of the cognitive load associated with the processing of different word orders. The question that must be addressed next is as follows: Why are grammatical functions so closely related to processing difficulties? Tamaoka et al. (2005) suggested that the linear ordering of grammatical functions consistently correlated with syntactic complexity in a wide range of constructions. In particular, the scrambled NS-S order seen in (4b), (5b) and (6b) is associated with more complex syntactic representations involving an empty category, as compared with the canonical S-NS order of (4a), (5a) and (6a). Thus, in addition to all the necessary steps involved in the processing of a canonically ordered sentence, processing its scrambled counterpart requires the creation and insertion of an empty category into the parsed structure, as well as linking it to the preceding nonsubject constituent. Tamaoka et al. contended that these additional steps in mental computation led to the increased reaction times observed for sentences (4b), (5b) and (6b) in their experiments.

## 2.2 Order of adjuncts

We saw in Section 2.1 that the scrambling of arguments incurs a processing cost in Japanese. Parallel results from the processing of orthographically and phonologi-

cally presented sentences have been reported for many other languages, including Dutch (Frazier and Flores d'Arcais 1989), Finnish (Kaiser and Trueswell 2004), German (Grewe et al. 2007; Rösler et al. 1998; Weyerts et al. 2002), Korean (Kim 2012), Russian (Sekerina 1997), and Sinhalese (Kanduboda and Tamaoka 2012), among many others. Almost all previous processing studies of scrambling, however, have focused on scrambling of arguments, and few have examined the processing effects of adjunct scrambling. Koizumi and Tamaoka (2006) analyzed whether or not scrambling of and/or across adjuncts also yields a higher processing load, using Japanese sentences with adverbs.

Japanese adverbs can be divided into (at least) three broad classes based on their syntactic distribution: (i) adverbs that are initially merged with a projection of a verb (i.e., VP adverbs), (ii) adverbs that are initially merged with a projection of a tense (i.e., TP adverbs), and (iii) adverbs that are initially merged with a projection of a modal (i.e., MP adverbs) (Koizumi 1993; Kimura 2004; cf. also Minami 1974; Nakau 1980; Noda 1984; Takubo 1987; Cinque 1999). The canonical positions of the three classes of adverbs are schematically shown in (7), where A represents an adverb.

(7) $[_{CP}$ $[_{MP}$ (MP-A) $[_{TP}$ (TP-A) Subj (TP-A) $[_{VP}$ (VP-A) Obj (VP-A) V] T] M] C]

VP adverbs include manner and resultative adverbs such as *hayaku* 'fast' and *konagonani* 'into pieces.' Their canonical positions within a VP are c-commanded by the negative morpheme in short negation sentences such as (8), where the negative morpheme occurs between a verb stem and tense morpheme. VP adverbs, therefore, tend to be the focus of negation; hence, (8) is interpreted as 'I ran not fast' (i.e., 'I ran slowly').

(8) $[_{VP}$ *hayaku* *hasir*]-*ana-katta*
fast run-NEG-PST
'(I) did not run fast.'

TP adverbs include time adverbs such as *kinoo* 'yesterday' and *kyoo* 'today.' Their canonical positions within a TP are outside the c-command domain of the negative morpheme in short negation sentences. Thus, in the short negation sentence (9a), the verb 'run' is negated rather than the adverb 'yesterday.' TP adverbs, however, can be the target of negation in long negation sentences with *wakedewanai*, which means 'it is not the case' and takes a TP as its complement. Therefore, the preferred reading of (9b) is 'the time when I ran was not yesterday.'

(9) a. $[_{TP}$ *kinoo* $[[_{VP}$ *hasir*]-*ana*]-*katta*]
yesterday run-NEG-PST
'(I) did not run yesterday.'

b. [$_{TP}$ *kinoo* [$_{VP}$ *hasir*]-*ta*] *wakedewanai*
yesterday run-PST it.is.not.the.case
'It is not the case that (I) ran yesterday.'

Finally, MP adverbs include various types of modal adverbs such as *osoraku* 'probably' and *saiwai* 'fortunately.' MP adverbs occur outside the c-command domain of the negative morpheme in both short and long negation sentences. Hence, they cannot be the target of negation, as shown in (10).

(10) a. [$_{MP}$ *osoraku* [$_{TP}$ [$_{VP}$ *hasir*]-*ana-katta*] *daroo*]
Probably run-NEG-PST seem
'Probably (he or she) did not run.'

b. [$_{MP}$ *osoraku* [[$_{TP}$ [$_{VP}$ *hasir*]-*ta*] *wakedewanai*] *daroo*]
probably run-PST it.is.not.the.case seem
'Probably it is not the case that (he or she) ran.'

According to the structure shown in (7), for sentences with a VP adverb, subject-adverb-object-verb (SAOV) and subject-object-adverb-verb (SOAV) are the canonical word orders, and adverb-subject-object-verb (ASOV) is a noncanonical derived word order involving adverb scrambling (11a), (11b) and (11c).

(11) a. ASOV (derived word order with a VP adverb)

| | | | | | |
|---|---|---|---|---|---|
| *Yukkuri* | *Taroo* | *ga* | *sinbun* | *o* | *yonda* |
| slowly | Taro | NOM | newspaper | ACC | read |

'Taro read a newspaper slowly.'

b. SAOV (canonical word order with a VP adverb)

| | | | | | |
|---|---|---|---|---|---|
| *Taroo* | *ga* | *yukkuri* | *sinbun* | *o* | *yonda* |
| Taro | NOM | slowly | newspaper | ACC | read |

c. SOAV (canonical word order with a VP adverb)

| | | | | | |
|---|---|---|---|---|---|
| *Taroo* | *ga* | *sinbun* | *o* | *yukkuri* | *yonda* |
| Taro | NOM | newspaper | ACC | slowly | read |

Similarly, for sentences with a TP adverb, ASOV and SAOV are the canonical word orders, and SOAV is a derived word order (12a), (12b) and (12c). Further, for sentences with an MP adverb, ASOV is the canonical word order, and SAOV and SOAV are noncanonical word orders (13a), (13b) and (13c).

(12) a. ASOV (canonical word order with a TP adverb)

| | | | | | |
|---|---|---|---|---|---|
| *Kinoo* | *Taroo* | *ga* | *kabin* | *o* | *kowasita* |
| yesterday | Taro | NOM | vase | ACC | broke |

'Taro broke a vase yesterday.'

b. SAOV (canonical word order with a TP adverb)

| | | | | | |
|---|---|---|---|---|---|
| *Taroo* | *ga* | *kinoo* | *kabin* | *o* | *kowasita* |
| Taro | NOM | yesterday | vase | ACC | broke |

c. SOAV (derived word order with a TP adverb)

| | | | | | |
|---|---|---|---|---|---|
| *Taroo* | *ga* | *kabin* | *o* | *kinoo* | *kowasita* |
| Taro | NOM | vase | ACC | yesterday | broke |

(13) a. ASOV (canonical word order with an MP adverb)

| | | | | | |
|---|---|---|---|---|---|
| *Zannennagara* | *Taroo* | *ga* | *syoosin* | *o* | *zitaisita* |
| unfortunately | Taro | NOM | promotion | ACC | refused |

'Unfortunately, Taro refused (an offer of) promotion.'

b. SAOV (derived word order with an MP adverb)

| | | | | | |
|---|---|---|---|---|---|
| *Taroo* | *ga* | zannennagara | *syoosin* | *o* | *zitaisita* |
| Taro | NOM | unfortunately | *promotion* | *ACC* | *refused* |

c. SOAV (derived word order with an MP adverb)

| | | | | | |
|---|---|---|---|---|---|
| *Taroo* | *ga* | *syoosin* | *o* | *zannennagara* | *zitaisita* |
| Taro | NOM | promotion | ACC | unfortunately | refused |

The relationship between the adverb classes and canonicity of word order is summarized in (14).

(14)

| Adverb Class | Canonical Word Order(s) | Derived Word Order(s) |
|---|---|---|
| VP adverbs: | SAOV and SOAV | ASOV |
| TP adverbs: | ASOV and SAOV | SOAV |
| MP adverbs: | ASOV | SAOV and SOAV |

In psycholinguistic literature, it is generally believed that, other things being equal, the human parser processes canonical word orders faster than derived word orders. If this generalization holds for the ordering of adjuncts as well, the analysis summarized in (14) predicts that in sentences with VP adverbs, SAOV and SOAV are processed faster than ASOV; in sentences with TP adverbs, ASOV and SAOV are processed faster than SOAV; and in sentences with MP adverbs, ASOV is processed faster than SAOV and SOAV. These predictions are summarized in (15).

(15) Predictions ('X < Y' stands for 'X is processed faster than Y.')
  a. VP adverbs: {SAOV, SOAV} < ASOV
  b. TP adverbs: {ASOV, SAOV} < SOAV
  c. MP adverbs: ASOV < {SAOV, SOAV}

Koizumi and Tamaoka (2006) tested these predictions by performing a reading experiment involving a sentence plausibility judgment task (Chujo 1983; Stromswold

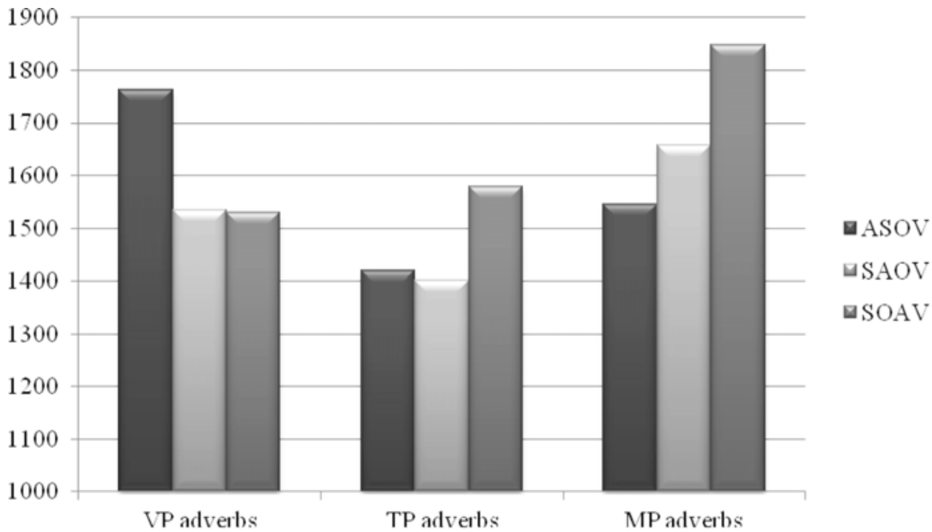



**Figure 1:** Reaction times (ms) for sentences with three classes of adverbs (Based on the data of Koizumi and Tamaoka 2006)

et al. 1996; Tamaoka et al. 2005). In this experiment, transitive sentences with adverbs such as those in (11)–(13), as well as semantically anomalous filler sentences, were visually presented to the participants, in the center of a computer screen, in random order. The participants were instructed to respond as quickly and accurately as possible in deciding whether the sentences were correct. They registered their responses by pressing either the "yes" or "no" button. The length of time between the appearance of a sentence on the screen and the button press was recorded as the response time.

The results of the experiment (Figure 1) confirmed all the predictions in (15). In sentences with VP adverbs, the response times were reliably longer for ASOV than for either SAOV or SOAV, while the response times for the latter two did not differ significantly. In sentences with TP adverbs, the response times were longer for SOAV than for ASOV and SAOV, with the response times for the latter two being comparable. In sentences with MP adverbs, ASOV was processed faster than SAOV, which, in turn, was processed faster than SOAV. These results taken together support the analysis of adverb distribution represented in (7); more crucially, they show that not only scrambling of and/or across arguments but also scrambling of and/or across adjuncts incurs a higher processing load.

## 2.3 Contextual effects

It has often been argued, as we have seen above, that noncanonical structures are inherently more difficult to process than canonical structures because they are

syntactically more complex and, hence, computationally more costly to represent than their canonical counterparts (e.g., De Vincenzi 1991; Frazier and Flores d'Arcais 1989; Tamaoka et al. 2005). Some researchers have also noted that the use of non-canonical structures is motivated by discourse-pragmatic factors. In particular, studies of sentence comprehension reveal that given-new ordered sentences (i.e., sentences in which given information is mentioned early and new information later) are read faster than all other sentences (Clifton and Frazier 2004; Haviland and Clark 1974; Ishida 1999). Furthermore, given-new ordering effects have been observed in sentence production, and many studies have reported that given-new ordered sentences are easier to remember and recall than other sentences (Bock and Irwin 1980; Bock and Warren 1985; Ferreira and Yoshita 2003; Vande Kopple 1982). Together, these studies show that given-new ordered sentences are easier to process.

Using Finnish, Kaiser and Trueswell (2004) extended these previous studies by examining interaction between the effects of syntactic structure (SS) (canonical SVO order vs. non-canonical OVS order) and those of information structure (IS) (given-new order vs. new-given order) on sentence processing load. They found that the given-new order was processed faster than the new-given order, but this IS effect was larger in non-canonical OVS sentences than in canonical SVO sentences. In particular, the interaction between SS and IS was observed at the second NP (O of SVO and S of OVS).

Given that an interaction of SS and IS was observed for Finnish, a head-initial language, Imamura and Koizumi (2008a) investigated whether or not a similar interaction also occurs in Japanese, a head-final language. Four types of two-sentence passages such as those in (16)–(19) were used for their experiment with a sentence plausibility judgment task. Each passage consisted of a context sentence and target sentence. The context sentences were all copula sentences, and the target sentences were all transitive sentences. NPs at the focus position in the context sentences (e.g., *Kuroki* in (16a)) were reused in the immediately following target sentences. These NPs denote given information in the target sentences, with the result that either the subject or object in the target sentences was given information. On the other hand, NPs that were not used in context sentences (e.g., *Kaneda* in (16b)) were new information in the target sentences. In sentences (16)–(19), as well as (22)–(25) below, boldface indicates an NP that represents old information.

(16) $[S_{given}\ O_{new}\ V]$

a. *Gaimusyoo* *no* *zikan* *wa* *Kuroki-da.*
the.Ministry.of.Foreign.Affairs GEN vice-minister TOP Kuroki-COP
'It is Kuroki who is the vice-minister of the Ministry of Foreign Affairs.'

b. ***Kuroki*** ***ga*** *Kaneda* *o* *mukaeta.*
Kuroki NOM Kaneda ACC welcomed
'Kuroki welcomed Kaneda.'

(17) [$S_{new}$ $O_{given}$ V]

a.

| | | | | |
|---|---|---|---|---|
| *Gaimusyoo* | *no* | *zikan* | *wa* | *Kaneda-da.* |
| the.Ministry.of.Foreign.Affairs | GEN | vice-minister | TOP | Kaneda-COP |

'It is Kaneda who is the vice-minister of the Ministry of Foreign Affairs.'

b.

| | | | | |
|---|---|---|---|---|
| *Kuroki* | *ga* | ***Kaneda*** | ***o*** | *mukaeta.* |
| Kuroki | NOM | Kaneda | ACC | welcomed |

'Kuroki welcomed Kaneda.'

(18) [$O_{given}$ $S_{new}$ V]

a.

| | | | | |
|---|---|---|---|---|
| *Gaimusyoo* | *no* | *zikan* | *wa* | *Kaneda-da.* |
| the.Ministry.of.Foreign.Affairs | GEN | vice-minister | TOP | Kaneda-COP |

'It is Kaneda who is the vice-minister of the Ministry of Foreign Affairs.'

b.

| | | | | |
|---|---|---|---|---|
| ***Kaneda*** | ***o*** | *Kuroki* | *ga* | *mukaeta.* |
| Kaneda | ACC | Kuroki | NOM | welcomed |

'Kuroki welcomed Kaneda.'

(19) [$O_{new}$ $S_{given}$ V]

a.

| | | | | |
|---|---|---|---|---|
| *Gaimusyoo* | *no* | *zikan* | *wa* | *Kuroki-da.* |
| the.Ministry.of.Foreign.Affairs | GEN | vice-minister | TOP | Kuroki-COP |

'It is Kuroki who is the vice-minister of the Ministry of Foreign Affairs.'

b.

| | | | | |
|---|---|---|---|---|
| *Kaneda* | *o* | ***Kuroki*** | ***ga*** | *mukaeta.* |
| Kaneda | ACC | Kuroki | NOM | welcomed |

'Kuroki welcomed Kaneda.'

The experiment had a 2 × 2 factorial design, with SS (canonical/scrambled) and IS (given-new/new-given) as the factors. Hence, there were four experimental conditions, as shown in (16)–(19).

Stimuli were presented to the participants in random order, in the center of the computer screen, one sentence at a time. Participants were instructed to respond as quickly and accurately as possible. They were asked to judge whether each sentence was semantically acceptable or unacceptable and indicate their response by pressing the left mouse button for "yes" and the right mouse button for "no." The reaction times were registered from the point of the stimulus presentation to the point when participants clicked the mouse to answer.

The results (Figure 2) show that the target sentences were processed significantly faster with the canonical word order than with the scrambled word order. This finding is in agreement with the results of previous studies (Chujo 1983; Koizumi and Tamaoka 2010; Mazuka, Ito and Kondo 2002; Miyamoto and Takahashi 2002; Tamaoka et al. 2005). In addition, the new-given word order took longer to process than the given-new word order. This finding also agrees with the results of

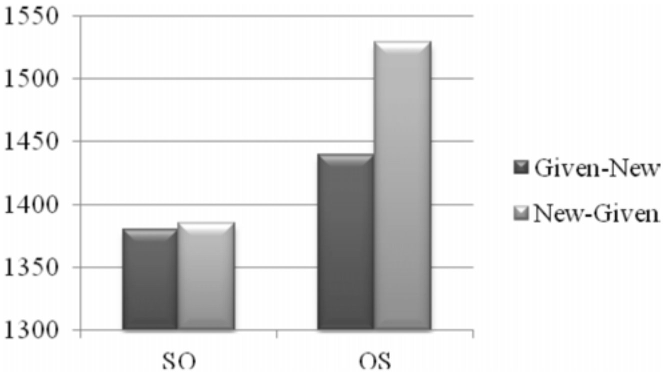



**Figure 2:** Reaction times (ms) for sentences in context (Adapted from Imamura and Koizumi 2008a)

previous studies (Haviland and Clark 1974; Kaiser and Trueswell 2004). More importantly, the interaction between SS and IS was significant in reading times because there is a significant difference between [$O_{given}$ $S_{new}$ V] and [$O_{new}$ $S_{given}$ V], but not between [$S_{given}$ $O_{new}$ V] and [$S_{new}$ $O_{given}$ V]. In other words, given-new status has an effect on processing load in marked but not unmarked word order. This coincides with previous linguistic studies that claim the marked pattern occurs only in the licensing context, whereas the unmarked pattern is contextually unrestricted (e.g., Kuno 1978; Aissen 1992). Thus, Imamura and Koizumi (2008a) revealed that the SS-IS interaction is not restricted to head-initial languages.

Imamura and Koizumi (2008a) presented a whole target sentence at a time; hence, it is not clear from their results at which point in time the interaction between SS and IS arises. Imamura and Koizumi (2008b) then examined the time course of the interaction between SS and IS in sentence comprehension in Japanese, a head-final language, from the perspective of both verb-driven and incremental processing models. The explanation based on the verb-driven processing model predicts that structure building should be delayed until a verb is reached (Pritchett 1988, 1991, 1992; Mulders 2002) and that the effect of IS on SS should arise at the verb (V in SOV and OSV) because, in a verb-driven model, this is the first point at which it is clear that a discourse violation occurs in scrambled word order. In contrast, incremental models anticipate that the parser processes sentences incrementally without waiting for crucial information from verbs (Kamide, Altmann, and Haywood 2003; Kutas and Hillyard 1980; Marslen-Wilson 1975) and that the SS-IS interaction should occur at the second NP (O in SOV and S in OSV), because it is the first point at which it is clear that a discourse violation arises if the parser incrementally processes sentences. Two hypotheses are listed as (20) and (21).

(20) Prediction by verb-driven models
If the Japanese parser delays structure building until a verb is encountered, then the interaction between SS and IS should arise at the third phrase (V).

(21) Prediction by incremental models
If the Japanese parser builds structure incrementally, then the interaction between SS and IS should arise at the second phrase (O in SOV or S in OSV), because this is the first point at which it is clear that a discourse violation occurs in scrambled word order.

It is important to emphasize that both hypotheses predict the interaction would arise at the third phrase (O in SVO and S in OVS) in SVO languages such as Finnish. Thus, it is necessary to examine this issue in an SOV language such as Japanese.

The stimuli used in Imamura and Koizumi's (2008b) experiment were the same as those used in Imamura and Koizumi (2008a), except that an auxiliary verb (*rasii* 'be likely' or *sooda* 'I hear that') was added to the end of each target sentence to avoid the possibility that end-of-sentence wrap-up effects would arise at the verb position, as shown in (22)–(25).

(22) [$S_{given}$ $O_{new}$ V]

a.

| P1 | | P2 | | P3 |
|---|---|---|---|---|
| *Gaimusyoo* | *no /* | *zikan* | *wa /* | *Kuroki-da. /* |
| the.Ministry.of.Foreign.Affairs | GEN | vice-minister | TOP | Kuroki-COP |

'It is Kuroki who is the vice-minister of the Ministry of Foreign Affairs.'

b.

| P4 | | P5 | | P6 | P7 |
|---|---|---|---|---|---|
| ***Kuroki*** | ***ga*** */* | *Kaneda* | *o /* | *mukaeta /* | *rashii. /* |
| Kuroki | NOM | Kaneda | ACC | welcomed | is.likely |

'It is likely that Kuroki welcomed Kaneda.'

(23) [$S_{new}$ $O_{given}$ V]

a.

| P1 | | P2 | | P3 |
|---|---|---|---|---|
| *Gaimusyoo* | *no /* | *zikan* | *wa /* | *Kaneda-da. /* |
| the.Ministry.of.Foreign.Affairs | GEN | vice-minister | TOP | Kaneda-COP |

'It is Kuroki who is the vice-minister of the Ministry of Foreign Affairs.'

b.

| P4 | | P5 | | P6 | P7 |
|---|---|---|---|---|---|
| *Kuroki* | *ga /* | ***Kaneda*** | ***o*** / | *mukaeta /* | *rasii. /* |
| Kuroki | NOM | Kaneda | ACC | welcomed | is.likely |

'Kuroki welcomed Kaneda.'

(24) [$O_{given}$ $S_{new}$ V]

a.

| P1 | | P2 | | P3 |
|---|---|---|---|---|
| *Gaimusyoo* | *no /* | *zikan* | *wa /* | *Kaneda-da. /* |
| the.Ministry.of.Foreign.Affairs | GEN | vice-minister | TOP | Kaneda-COP |

'It is Kaneda who is the vice-minister of the Ministry of Foreign Affairs.'

| | | | | | | |
|---|---|---|---|---|---|---|
| b. | P4 | | P5 | | P6 | P7 |
| | ***Kaneda*** | ***o*** / | *Kuroki* | *ga* / | *mukaeta* / | *rasii.* / |
| | Kaneda | ACC | Kuroki | NOM | welcomed | is.likely |
| | 'Kuroki welcomed Kaneda.' | | | | | |

(25) [$O_{new}$ $S_{given}$ V]

| | | | | | | |
|---|---|---|---|---|---|---|
| a. | P1 | | P2 | | P3 | |
| | *Gaimusyoo* | *no* / | *zikan* | *wa* / | *Kuroki-da.* / | |
| | the.Ministry.of.Foreign.Affairs | GEN | vice-minister | TOP | Kuroki-COP | |
| | 'It is Kuroki who is the vice-minister of the Ministry of Foreign Affairs.' | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| b. | P4 | | P5 | | P6 | P7 |
| | *Kaneda* | *o* / | ***Kuroki*** | ***ga*** / | *mukaeta* / | *rasii.* / |
| | Kaneda | ACC | Kuroki | NOM | welcomed | is.likely |
| | 'Kuroki welcomed Kaneda.' | | | | | |

Participants were instructed to read the stimuli phrase by phrase at their own pace (self-paced reading task). They were timed in a phrase-by-phrase non-cumulative moving-window reading task. After they finished reading each two-sentence passage, the participants were asked to push the "yes" button if the two-sentence passage they had just read made sense and the "no" button if it was semantically anomalous (plausibility judgment task). They were instructed to judge as accurately and as quickly as possible. The phrase by phrase reading times and error rates were measured.

Error rates were lower for the given-new word order sentences than for the new-given word order sentences ([$S_{given}$ $O_{new}$ V] = 2.74%, [$S_{new}$ $O_{given}$ V] = 5.65%, [$O_{given}$ $S_{new}$ V] = 3.08%, [$O_{new}$ $S_{given}$ V] = 4.01%). In regard to reading times, it was found that P4 (S in SOV and O in OSV) was read significantly faster in the given-new word order than in the new-given word order (Figure 3). No other main effects or interactions were significant. P5 (O in SOV and S in OSV) was read significantly faster in the canonical word order than in the scrambled word order, and it was read significantly faster in the given-new word order than in the new-given word order. The interaction between these factors was significant: The difference between [$O_{given}$ $S_{new}$ V] and [$O_{new}$ $S_{given}$ V] was larger than that between [$S_{given}$ $O_{new}$ V] and [$S_{new}$ $O_{given}$ V] in reading times. P6 (verbs) was read significantly faster in the canonical word order than in the scrambled word order. No other effects or interactions were significant. Finally, at the auxiliary verbs (P7), none of the main effects or interactions were significant.

Important to our concern is that the SS-IS interaction was observed at P5 (O in SOV and S in OSV). This means that, in reading times, the difference between [$O_{given}$ $S_{new}$ V] and [$O_{new}$ $S_{given}$ V] is larger than that between [$S_{given}$ $O_{new}$ V] and [$S_{new}$ $O_{given}$ V]. If the parser processes head-final languages incrementally, P5 is the first point

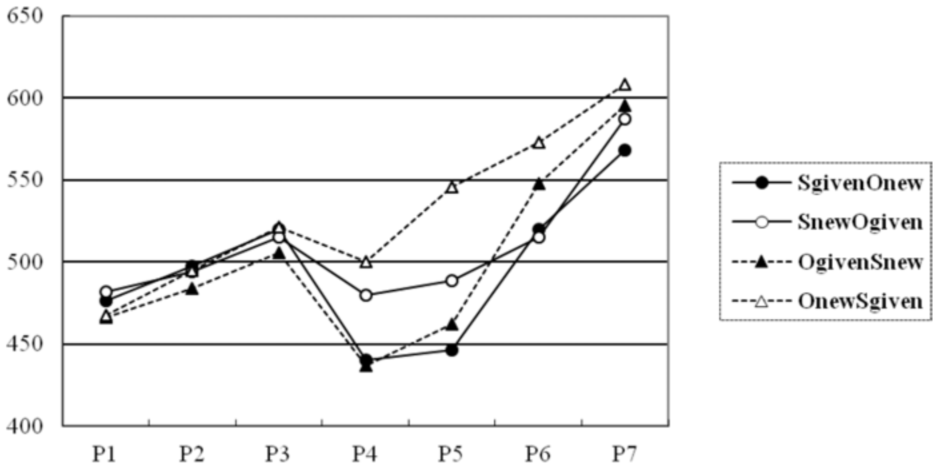



**Figure 3:** Reaction times (ms) for self-paced reading task (Adapted from Imamura and Koizumi 2008b)

to indicate that a discourse violation arises in scrambled ordering. As mentioned above, the scrambled word order occurs only in the licensing context, while the canonical word order is contextually unrestricted, and hence, the interaction occurs. Thus, prediction (21) was borne out, which supports an incremental model in that there is no delay in parsing, and NPs are associated with each other immediately according to the given-new status available.

We have reviewed two reading experiments on the processing of word order variants in a head-final language. These experiments investigated two issues: The first experiment concentrated on whether SS-IS interaction is confined to head-initial languages; the second examined the time course of the interaction between SS and IS to determine whether the parser processes head-final languages based on a verb-driven or incremental processing model. The former indicates that SS-IS interaction is not restricted to head-initial languages. Analysis of the effect of IS on SS showed that canonical word order is not sensitive to given-new status, but that given-new status influences the processing of scrambled word order. These results are consistent with those of previous theoretical linguistic studies, which report that the marked pattern occurs only in the licensing context, whereas the unmarked pattern is contextually unrestricted. In the latter experiment, the SS-IS interaction was observed at the second NP (O in SOV or S in OSV). This result supports the incremental model, in which the NPs must be associated without delay, because the point where the interaction arose is the first point where it is clear that a discourse violation occurs in scrambled word order, if the parser incrementally constructs the structure. The interaction between SS and IS was observed because canonical ordering was less sensitive to given-new status than scrambled ordering was. This result

accords with Kuno (1978), who insisted that sentences involving marked violations of discourse principles are unacceptable, but those that involve unmarked violations of discourse principles go unpenalized and are acceptable. However, some qualification is in order. That is, the results of the second experiment clearly demonstrate that [$S_{new}$ $O_{given}$ V] is penalized for violating the given-new ordering in terms of error rates as well as reaction times. It may not be reasonable to say that the unmarked pattern is unpenalized for violation of discourse rules because given-new status has little influence on the canonical condition. Thus, Imamura and Koizumi (2008b) suggested that sentences that involve unmarked violations of discourse principles are less penalized and more acceptable than those that involve marked violations.

# 3 From processing to grammar

In this section, we examine sentence processing studies belonging to category (II), i.e., the experimental syntax in a narrow sense. As with the sentence processing studies in category (I), studies in (II) also use experimental techniques. However the primary focus of the studies in category (II) is on evaluating linguistic theories/hypotheses rather than processing theories/hypotheses.

## 3.1 Cartography and phrase structural status of tense and aspect

A research program widely known as Cartography, or the cartographic approach to syntactic structure, attempts to draw maps of syntactic configurations with as much precision and detail as possible (Cinque and Rizzi 2008; see also Cinque 2002; Rizzi 2004; Belletti 2004). The strongest position of the cartographic approach assumes that the distinct hierarchies of functional projections may be universal in the type of heads and specifiers that they involve, in their number, and in their relative order (Rizzi 1997; Cinque 1999, 2002). A weaker position would assume that languages may differ in the type or number of functional projections they select from a universal inventory or in their order (Fukui 1995; Thráinsson 1996). According to the strongest position, tense and aspect, for example, separately project their own maximal projections, that is, TP and AspP, respectively, in every language, as there is positive evidence that they do in some languages. In contrast, the weaker position might accept that some languages lack one or both of TP and AspP. Kimura, Kim and Koizumi (2005) presented an empirical argument to the effect that the strongest position is not tenable.

Generally, aspectual adjuncts such as *sibasiba* ‘frequently’ follow tense adjuncts but precede manner/resultative adjuncts when they co-occur. Thus, the sequential order of tense-aspectual-manner/resultative adjuncts cannot be inverted, as shown by (26) and (27).

(26) a.

| | | | | | | |
|---|---|---|---|---|---|---|
| *Mary* | *ga* | *konsyuu* | *sibasiba* | *tosyokan* | *o* | *tukatta.* |
| Mary | NOM | this.week | frequently | library | ACC | used |

‘Mary frequently used the library this week.’

b.

| | | | | | | |
|---|---|---|---|---|---|---|
| ??*Mary* | *ga* | *sibasiba* | *konsyuu* | *tosyokan* | *o* | *tukatta.* |
| Mary | NOM | frequently | this.week | library | ACC | used |

(27) a.

| | | | | | | |
|---|---|---|---|---|---|---|
| *Mary* | *ga* | *tabitabi* | *yukkuri* | *tosyokan* | *o* | *tukatta.* |
| Mary | NOM | frequently | leisurely | library | ACC | used |

‘Mary frequently took her time in using the library.’

b.

| | | | | | | |
|---|---|---|---|---|---|---|
| **Mary* | *ga* | *yukkuri* | *tabitabi* | *tosyokan* | *o* | *tukatta.* |
| Mary | NOM | leisurely | frequently | library | ACC | used |

The permuted word order of a tense adjunct following an aspectual adjunct as in (26b) is seemingly worse than its canonical counterpart (26a). Likewise in (27a), the order of an aspectual adjunct preceding a manner adjunct cannot be altered into the manner-aspect order, as is credible in (27b). This may well be sufficient evidence that aspectual adjuncts exist as a unique category, apart from manner or tense adjuncts.

We saw in Section 2.2 that VP manner adjuncts are base-generated in a position before or after the object, and tense adjuncts on either side of the subject. Observations of aspectual adjuncts so far suggest that their canonical position is neither one, but rather, between the subject and object, on the borderline of TP and VP. However, it is also essential to confirm which phrase they belong to. Kimura, Kim and Koizumi (2005) adopted an online sentence plausibility judgment task, using a similar method to that of Koizumi and Tamaoka (2006), to verify the canonical position of aspectual adjuncts and word order of Japanese, which were not taken into account in the experiment of Koizumi and Tamaoka (2006).

To deal with the issues concerning Japanese aspectual adjuncts and their base position, Kimura, Kim and Koizumi (2005) postulated two possible constructions for sentences that include aspectual adjuncts. One is to adopt an Aspectual Phrase analysis (cf. Travis 2005; McClure 1994; Cowper 1999; Ernst 2002; Dubinsky and Hamano 2003, among many others) and to assume that a maximal projection AspP exists between TP and VP in Japanese, attaching the aspectual adjuncts within that projection. The configurationality of the AspP node and possible structure of a sentence that includes aspectual adjuncts are shown in (28).

(28) Hypothesis 1: AspP Analysis

a. $[_{\mathrm{TP}}$ (TP-A) S (TP-A) $[_{\mathrm{AspP}}$ (AspP-A) $[_{\mathrm{VP}}$ (VP-A) O (VP-A) V] ASP] T]

b.

| | | | | | | |
|---|---|---|---|---|---|---|
| *Taroo* | *ga* | *konsyuu* | *tabitabi* | *hon* | *o* | *yonda.* |
| Taro | NOM | this.week | frequently | book | ACC | read |

‘Taro frequently read books this week.’

c. $[_{\mathrm{TP}}$ *Taroo ga (konsyuu)* $[_{\mathrm{AspP}}$ *(tabitabi)* $[_{\mathrm{VP}}$ *hon o yonda*] ASP] T]

Notice that various types of adjuncts are attached to different projections in (28), essentially following the spirit of Chinque (1999) and the strongest version of the cartographic approach. This structure naturally provides an explanation for the sequential order of tense-aspect-VP adjuncts.

In spite of the AspP analysis assumed in (28), there still is a possibility that aspectual adjuncts exist inside the same projection as tense adjuncts, given that tense and aspectual adjuncts are semantically closely related to each other. This hypothesis of denying the existence of AspP for aspectual adjuncts in Japanese, which we refer to as the Inner IP Aspect analysis, can be diagrammatically configured as in (29):

(29) Hypothesis 2: Inner IP Aspect Analysis
a. [$_{IP}$ (TP-A/Asp-P) S (TP-A/AspP-A) [$_{VP}$ (VP-A) O (VP-A) V] I]
b. [$_{IP}$ *Taroo ga (konsyuu) (tabitabi)* [$_{VP}$ *hon o yonda*] I]

Based on this structure, it is fair to say that tense adjuncts and aspectual adjuncts can appear on either side of a subject, keeping their sequential order even when they co-occur in a single sentence. The Inner IP Aspect analysis is consistent with the weaker position of the cartographic approach to syntactic structure but not with the strongest position.

The two analyses above make different predictions for the participants' reaction times in the processing experiment. If we bear in mind that Japanese transitive constructions allow three potential word orders, including aspectual adjuncts, the AspP analysis predicts that ASOV or SOAV word order takes a longer time to process than SAOV order, while the Inner IP Aspect analysis predicts that SOAV takes a longer reading time than both ASOV and SAOV.

(30) Predictions for Reaction Times
a. AspP analysis: SAOV < ASOV/SOAV
b. Inner IP Aspect analysis: ASOV/SAOV < SOAV

The goal of Kimura, Kim and Koizumi's (2012) experiment was to determine the canonical position of aspectual adjuncts in Japanese. If either one of the two hypotheses given above is correct, then we can draw a conclusion as to which phrase the aspectual adjuncts belong to in the phrase structure.

Sentences with aspectual adjuncts in three different orders (ASOV, SAOV, and SOAV), such as those in (31), were used as stimuli in their experiment.

(31) a. *Itumo Taroo ga hon o yondeiru.*
always Taro NOM book ACC be.reading
'Taro is always reading a book.'

| | | | | | | |
|---|---|---|---|---|---|---|
| b. | *Taroo* | *ga* | *itumo* | *hon* | *o* | *yondeiru.* |
| | Taro | NOM | always | book | ACC | be.reading |

| | | | | | | |
|---|---|---|---|---|---|---|
| c. | *Taroo* | *ga* | *hon* | *o* | *itumo* | *yondeiru.* |
| | Taro | NOM | book | ACC | always | be.reading |

The stimulus sentences were randomly presented in the center of the screen. A sentence appeared in its entirety on the screen, and participants were asked to look at it and decide whether it was semantically plausible or not, by clicking the left mouse button for "yes" and the right button for "no." Participants' responses were timed from the point of stimulus presentation to the point when they clicked the mouse to answer.

The results of the experiment provide evidence that native speakers of Japanese process ASOV (1814 ms) and SAOV (1804 ms) word order faster than SOAV (1991 ms), as far as aspectual adjuncts are concerned. This finding supports Hypothesis 2, the Inner IP Aspect analysis. The supported analysis excludes the existence of AspP in Japanese and posits the aspectual adjuncts inside the same maximal projection as tense adjuncts, which is called IP in Kimura, Kim and Koizumi (2005). The significantly longer reading time required for SOAV implies that the position between object and verb is not a canonical position for aspectual adjuncts at all. The results of the current experiment and of Koizumi and Tamaoka (2006) support Kimura, Kim and Koizumi's (2005) proposal that the object undergoes scrambling across the aspectual adjunct in the derivation of SOAV sentences, which can serve to explain the longer reading times due to the extra processing load.

A problem that still needs to be explained is the naturalness of the co-existing adjunct order, tense to aspect. Consider again the examples in (26). If we conclude that both the tense adjunct *konsyuu* and aspectual adjunct *sibasiba* co-occur inside IP, why is their order restricted as in (26), instead of being freely permutable within the phrase? The reason for this phenomenon is that tense adjuncts in Japanese must take a wider scope than aspectual adjuncts do. Thus, the order is always constrained in that tense adjuncts appear closer to the sentence-initial position, and when both adjunct types appear together, aspectual adjuncts occur within the scope of tense adjuncts. This hypothesis makes sense if we consider that most languages do not have an aspectual feature outside (or, perhaps "higher" than) the tense feature. Tense refers to a certain point or length of time, but aspect can only internally indicate such a point's progressivity or perfectivity within the range of time defined by tense. Therefore, Japanese is as likely as other languages to have an aspectual feature whose interpretation is dependent on tense.

In conclusion, Kimura, Kim and Koizumi (2005) have shown that tense and aspectual adjuncts in Japanese are both base-generated in the same phrase, which might be called IP, rather than merged to different phrases, that is, TP and AspP,

respectively. This finding effectively denies the existence of separate maximal projections for tense and aspect in Japanese, constituting potentially strong counter-evidence against the strongest position of the cartographic approach to syntactic structure in general and to the universal hierarchy of phrase structure and adverb distribution proposed by Chinque (1999) in particular.

## 3.2 VP-internal subject position and the Extended Projection Principle

Traditionally, the subject was defined as an NP immediately dominated by an S node (Chomsky 1957). Thus, when Saito and Hoji (1983) argued that in Japanese, the object is base-generated within the VP (32a) and that when it occurs to the left of the subject, it has undergone scrambling (32b), it was presumed that the subject is in a position directly under the S node throughout the derivation (see also Saito 1985 and Hoji 1985).

(32) a. $[_S$ S $[_{VP}$ O V]]

b. $[_S$ $O_i$ $[_S$ S $[_{VP}$ $t_i$ V]]

However, according to the Internal Subject Hypothesis, the "base" or "thematic" position of the subject (i.e., the external argument), as well as that of the object (i.e., the internal argument), is within the VP (Fukui 1986; Kitagawa 1986; Koopman and Sportiche 1988; Kuroda 1988), and when the subject is outside the VP, it has moved from its base position for some reason. For example, the Extended Projection Principle (EPP) requires that every clause have a subject occupying the canonical subject position, that is, that Spec TP be filled with a nominal element (Chomsky 1981). In research on Japanese, it is now a standard analysis that the subject in a canonically ordered SOV sentence moves from its thematic position to its derived position, Spec TP, as shown in (33a) or (33b), where a more recent proposal of a VP/vP distinction is also adopted (Miyagawa 1998; Kishimoto 2001).

(33) a. $[_{TP}$ $S_i$ $[_{VP}$ $t_i$ OV]]

b. $[_{TP}$ $S_i$ $[_{vP}$ $t_i$ $[_{VP}$ OV]]]

The discussions in the previous subsections were also based on this assumption, although the presence of a subject trace within the VP/vP was not explicitly mentioned because it was not relevant.

For Japanese OSV sentences, there are at least two competing analyses with respect to the placement of the subject. One is that the subject in OSV sentences, similar to the subject in SOV sentences, obligatorily moves to Spec TP, and the object moves to an even higher position, as shown in (34) (cf. Saito 2003, etc.).

(34) Analysis 1: $[_{TP}$ $O_j$ $S_i$ $[_{vP}$ $t_i$ $[_{VP}$ $t_j$ V]]]

Although this structure contains a vP-internal subject trace, it is fairly similar to the traditional structure presented in (32b) above. Therefore, it can be considered an updated version of the traditional structure.

A more innovative analysis was proposed by Miyagawa (2001), according to which it is possible that the subject stays at its base position within the vP, and only the object moves to Spec TP, as shown in (35).

(35) Analysis 2: $[_{TP}$ $O_i$ $[_{vP}$ $t_i$' S $[_{VP}$ $t_i$ V]]]

Note that the derivation of this structure involves two movements of the object: First, the object moves to the edge of vP, which is necessary for locality reasons, and second, the object moves to Spec TP (Miyagawa 2001; Miyagawa and Arikawa 2007). Hence, the OSV has a more costly derivation than SOV word order.

Part of the evidence for the proposal that the subject may stay in the vP, as in (35), comes from the scope interpretation. Recall that constituents within a vP (= former VP) are c-commanded by the negation in short negation sentences; however, this is not the case with constituents that belong to a TP or MP (Section 2.2). Thus, if the universal quantifier *zen'in* 'all (members)' occurs in the object position of SOV sentences, it may be interpreted as being inside the scope of negation, thereby yielding a partial negation reading, that is, 'not all.' If *zen'in* occurs in the subject position, it is interpreted as being outside the scope of negation.

(36) a. *Taroo ga zen'in o sikar-anakat-ta*
Taro NOM all ACC scold-NEG-PST
'Taro did not scold all.' (not > all, all > not)

b. *Zen'in ga Taroo o sikar-anakat-ta*
all NOM Taro ACC scold-NEG-PST
'All did not scold Taro.' (??not > all, all > not)

Significantly, Miyagawa (2001) observed that if the object scrambles across the subject *zen'in*, partial negation becomes easier to obtain with appropriate prosody, as exemplified in (37) (see also Miyagawa 2006 and Miyagawa and Arikawa 2007).

(37) *Sono tesuto* $o_i$ *zen'in ga* $t_i$ *uke-nakat-ta* (*yo/to omou*)
that test $ACC_i$ all NOM $t_i$ take-NEG-PST
'That test, all didn't take.'
(Miyagawa 2001: 299)

Miyagawa (2001) argued that the partial negation interpretation of (37) is readily explained if we assume that its subject occupies a vP-internal position (as in (35)) that is c-commanded by the negation.

The two competing analyses in (34) and (35) create different predictions for the processing of OSV sentences with VP adverbs in three different positions, such as those in (38).

(38) a. AOSV: *Yukkuri sinbun o Taroo ga yonda.*
slowly newspaper ACC Taro NOM read
'Taro read a newspaper slowly.'

b. OASV: *Sinbun o yukkuri Taroo ga yonda.*
newspaper.ACC slowly Taro NOM read

c. OSAV: *Sinbun o Taroo ga yukkuri yonda.*
newspaper.ACC Taro NOM slowly read

Before discussing the predictions, however, it is necessary to make Assumption 2 more precise. We have seen above that VP adverbs may be base generated to either the left or right of the base position of the object within VP. It is not yet clear, at this point, whether they can be base generated within vP, as shown in (39a), or if they cannot, as shown in (39b).

(39) Base positions of VP-adverbs: Two versions of Assumption 2
a. Assumption 2a: $[_{vP}$ (A) S (A) $[_{VP}$ (A) O (A) V]]
b. Assumption 2b: $[_{vP}$ S $[_{VP}$ (A) O (A) V]]

We will therefore consider both possibilities in the following discussion.

Let us now turn to the predictions of Analysis 1 and Analysis 2 for the processing of OSV sentences with VP adverbs. In Analysis 1, with either Assumption 2a or Assumption 2b, the VP adverb *yukkuri* 'slowly' occupies its base position within the VP/vP in the OSAV construction (38c) and has undergone scrambling in the other two orders (38a, b). This is schematically represented in (40). (The traces of the arguments are omitted, and ($t_i$) stands for the trace of the adverb that would be left within the VP under Assumption 2b.)

(40) Schematic structures of the sentences in (38) in Analysis 1:
a. $[_{TP}$ $A_i$ O S $[_{vP}$ … $t_i$' …$[_{VP}$ … ($t_i$) … V]]]
b. $[_{TP}$ O $A_i$ S $[_{vP}$ … $t_i$' …$[_{VP}$ … ($t_i$) … V]]]
c. $[_{TP}$ O S $[_{vP}$ … A … V]]

Thus, Analysis 1 predicts that AOSV and OASV are more difficult to process than OSAV.

Analysis 2 makes different predictions with Assumption 2a and Assumption 2b. With Assumption 2a, both OASV (38b) and OSAV (38c) are canonical word orders with respect to the adverb placement, and AOSV (38a) alone involves adverb scrambling, as shown in (41).

(41) Schematic structures of the sentences in (38) in Analysis 2 with Assumption 2a:
  a. $[_{TP}\ A_i\ O\ [_{vP} \ldots t_i\ S\ [_{VP} \ldots\ldots V]]]$
  b. $[_{TP}\ O\ [_{vP} \ldots AS\ [_{VP} \ldots V]]]$
  c. $[_{TP}\ O\ [_{vP} \ldots SA\ [_{VP} \ldots V]]]$

It is then expected that AOSV is more difficult to process than both OASV and OSAV.

With Assumption 2b, in contrast, AOSV (38a) involves two movements of the adverb: a movement to the edge of vP and a movement to Spec TP from the vP edge. OASV (38b) involves a movement of the adverb across the subject within vP. OSAV (38c) is a canonical word order with respect to the adverb. This is schematically shown in (42), which predicts that AOSV is more difficult to process than OASV, which in turn is more demanding than OSAV.

(42) Schematic structures of the sentences in (38) in Analysis 2 with Assumption 2b:
  a. $[_{TP}\ A_i\ O\ [_{vP}\ t_i' \ldots S\ [_{VP}\ t_i \ldots V]]]$
  b. $[_{TP}\ O\ [_{vP} \ldots A_i\ S\ [_{VP} \ldots t_i \ldots V]]]$
  c. $[_{TP}\ O\ [_{vP} \ldots S\ [_{VP} \ldots A \ldots V]]]$

The predictions regarding the cognitive load associated with the processing of sentences like (38) are summarized in (43).

(43) Predicted processing load

| | | |
|---|---|---|
| a. | Analysis 1 with Assumption 2a or 2b: | AOSV = OASV > OSAV |
| b. | Analysis 2 with Assumption 2a: | AOSV > OASV = OSAV |
| c. | Analysis 2 with Assumption 2b: | AOSV > OASV > OSAV |

Koizumi and Tamaoka (2010) tested the predictions in (43) by conducting a psycholinguistic experiment with a sentence plausibility judgment task that involved whole sentence reading (cf. Chujo 1983; Tamaoka et al. 2005). Semantically plausible triplets such as those in (38) were constructed using VP adverbs. The stimuli were presented to the participants in random order in the center of a computer screen,

one sentence at a time. The participants were instructed to judge whether or not the sentences made sense and report their response as quickly and accurately as possible by pressing either a "yes" or a "no" button. The duration between the stimulus presentation and the button press was recorded as the response time.

Overall, transitive sentences in which the object precedes the subject took longer to process in AOSV (1695 ms) than in OASV (1550 ms) or OSAV (1590 ms), and the processing times for the latter two were comparable, as indicated in (44).

(44) Overall reaction times
AOSV > OASV = OSAV

The results of the experiment summarized in (44) are consistent with the prediction of Analysis 2 with Assumption 2a (i.e., (43b)), but not with the prediction of Analysis 2 with Assumption 2b (43c) or that of Analysis 1 (43a). Most importantly, OASV word order was significantly less difficult to process than AOSV word order, contrary to the prediction of Analysis 1. This suggests that the subjects of Japanese transitive sentences may stay in the vP when they follow the objects, as has been argued and defended in a series of papers by Miyagawa (e.g., Miyagawa 2001, 2003).[3] This, in turn, supports the central premise of the Internal Subject Hypothesis, that is, the base position of the external argument is within the vP rather than outside it to begin with. Furthermore, the comparable reaction times for OASV and OSAV are consistent with Assumption 2a but not with Assumption 2b, indicating that VP adverbs can be initially merged to the left of the base position of the subject within the vP.

Thus far, we have assumed that the subject always moves to Spec TP in canonically ordered SOV sentences. However, there exists a differing view. In some versions of the Internal Subject Hypothesis, the subject in Japanese stays within the VP throughout the derivation, regardless of whether it precedes or follows the object (Fukui 1995; Kuroda 1988). Let us consider whether this proposal can account for the experimental results. With respect to this particular version of the Internal Subject Hypothesis, the discussion in Section 2.2 needs to be reinterpreted in such a way that VP adverbs initially occur in "the lower part of VP" and TP adverbs, in "the higher part of VP," as shown in (45).

(45) $[_{VP}$ (TP-A) S (TP-A) $[_{V'}$ (VP-A) O (VP-A) V]]

When the object undergoes scrambling as shown in (46), the subject stays in the same position as the subject in (45).

3 This by no means entails that the structure in (34), in which the subject as well as the object has moved out of the vP, is impossible to attain. On the contrary, there is good reason to believe that Japanese grammar allows not only the structures in (35) but also that in (34). If so, the argument in the text still holds true. See Koizumi and Tamaoka (2010) for details.

(46) $O_i$ S [$_{V'}$ (VP-A) $t_i$ (VP-A) V]

We shall refer to this analysis as Analysis 3. In Analysis 3, the VP adverb occupies its canonical position in OSAV sentences, such as (38c), and the adverb has undergone scrambling in AOSV sentences, such as (38a), as well as in OASV sentences, such as (38b), as indicated in (47). (The object trace is not represented.)

(47) Schematic structures of the sentences in (38) in Analysis 3
a. $A_i$ O S [$_{V'}$ . . . $t_i$ . . . V]]
b. O $A_i$ S [$_{V'}$ . . . $t_i$ . . . V]]
c. O S [$_{V'}$ . . . A . . . V]]

Because (47a) and (47b) involve adverb scrambling and are therefore more syntactically complex than (47c), it is predicted that sentences with a VP adverb are more difficult to process in AOSV and OASV than in OSAV.

(48) Predicted processing load:
Analysis 3: AOSV = OASV > OSAV

The prediction of Analysis 3 shown in (48) is the same as the prediction of Analysis 1 shown in (43a), and it is crucially incompatible with the results of the experiment reported above. Koizumi and Tamaoka (2010) concluded, therefore, that Analysis 3 cannot account for the distribution of Japanese adverbs and their processing data. In other words, the subjects of transitive sentences in Japanese must move to Spec TP when they precede objects, suggesting that Japanese T has the EPP feature.

To summarize, Koizumi and Tamaoka (2010) presented processing evidence for the hypothesis that the subject of a transitive verb in Japanese overtly moves to Spec TP when it precedes the object, but it may stay in situ within the vP when it follows the object. This, in turn, supports the central premise of the ISH, which states that the base position of the external argument is within the vP.

## 3.3 Syntactically basic word order

We have seen that experimental methods are quite useful in revealing both the nature of the human parser and the nature of the cognitive system of language. So far, we have restricted inquiry to Japanese. As a way of demonstrating that similar approaches can be beneficial in studies of other languages, we will briefly look at a case study of Kaqchikel, a language typologically different from Japanese.

Kaqchikel is one of the Mayan languages spoken in Guatemala. It is recognized as having verb-object-subject (VOS) as its basic word order, similar to many of the

other Mayan languages. In reality, however, the SVO word order is more frequently used than VOS, which comes in second by comparison. For this reason, Kaqchikel is often referred to as a language that is possibly shifting, or has already shifted, from VOS to SVO. Koizumi et al. (2014) conducted a sentence processing experiment to investigate whether Kaqchikel's syntactically determined basic word order is VOS or SVO. This resulted in a traditional hypothesis that the syntactically determined basic word order is VOS, even for contemporary native Kaqchikel speakers (see also Kiyama et al. 2013).

Like other Mayan languages, Kaqchikel is head-marking: Subjects and objects are unmarked, and the verb carries person-number agreement markers for both the object and subject. Kaqchikel is also an ergative language; it has an absolutive agreement marker for both the transitive verb object and the intransitive verb subject, and an ergative agreement marker for the transitive verb subject. As shown in (49), the verbal complex (predicate) begins with a compounding aspect morpheme that expresses aspect, tense, and modality, arranged in the sequence of [aspect-absolutive-ergative-verb stem].

(49) *Y-e'-in-to'*.
IC-ABS3PL-ERG1SG-help
'I help them.'

Since Kaqchikel is a pro-drop language, (49) functions as both independent speech and an independent sentence.

Although, like many other Mayan languages, Kaqchikel allows different grammatical word orders, its standard word order is "verb initial." If the sentence is irreversible as in case (50) (where the meaning of the sentence collapses with the reversal of object and subject), it can be interpreted in either VOS or VSO order. However, VOS is preferred.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (50) | a. | *X-Ø-u-chöy* | *ri* | *chäj* | *ri* | *ajanel.* | [VOS] |
| | | CP-ABS3SG-ERG3SG-cut | DET | pine.tree | DET | carpenter | |
| | b. | *X-Ø-u-chöy* | *ri* | *ajanel* | *ri* | *chäj.* | [VSO] |
| | | CP-ABS3SG-ERG3SG-cut | DET | carpenter | DET | pine.tree | |

'The carpenter cut the pine tree.'

In cases like (51 a) and (51b), where the sentence is semantically reversible (it makes sense when the object and subject are reversed), a VOS interpretation is overwhelmingly favored (even though a VSO interpretation is still possible).

| | | | | | | |
|---|---|---|---|---|---|---|
| (51) | a. | *X-Ø-r-oqotaj* | *ri* | *me's* | *ri* | *tz'i'.* |
| | | CP-ABS3SG-ERG3SG-run.after | DET | cat | DET | dog |

'The dog ran after the cat.'

b. *X-Ø-r-oqotaj* *ri* *tz'i'* *ri* *me's.*
CP-ABS3SG-ERG3SG-run.after DET dog DET cat
'The cat ran after the dog.'

In cases like (52), where the subject is preposed before the verb, the subject is naturally interpreted as topical or focused.

(52) *Ri* *ajanel* *x-Ø-u-chöy* *ri* *chäj.* [SVO]
DET carpenter CP-ABS3SG-ERG3SG-cut DET pine.tree
'The carpenter cut the pine tree.'

In this sense, SVO order is pragmatically marked. Furthermore, SVO word orders have traditionally required transformation of the predicate, such as by adding an agent-focus morpheme. Thus, it can be said that SVO is also marked from a morphological perspective. However, in modern Kaqchikel, it is possible to attain the SVO word order without transforming the morphological form of the verbal complex (retaining the same morphological form as in VOS or VSO), as is evident in example (52).

Based on the reasons above, in addition to the fact that historical evidence, such as stelae (stone monuments), suggests VOS as the basic word order of ancient Mayan, the majority of Mayan language researchers consider the syntactically determined basic word order of modern Kaqchikel to be VOS (Rodríguez Guaján 1994: 200; García Matzar and Rodríguez Guaján 1997: 333; Tichoc Cumes et al. 2000: 195; Ajsivinac Sian et al. 2004: 162). According to England (1991: 480), sentences with the abovementioned three types of word orders have the syntactic structures shown in (53) (see also Aissen 1992, Tada 1993, Coon 2010, and Preminger 2011).

(53) a. [VOS]
b. [[V $t_i$ S] $O_i$]
c. [$S_i$ [VO $t_i$]]

However, in terms of usage frequency, VOS order is only second, as SVO is used considerably more frequently (Maxwell and Little 2006; Kubo et al. 2012). Because conversations develop as a chain of topic, SVO word order with a topicalized subject seems to appear frequently (England 1991; Broady 1984; Skopeteas and Verhoeven 2009). Furthermore, some researchers argue for the possibility of an influence from Spanish, the most popular language spoken in Guatemala (cf. England 1991: 475). For these reasons, wherein SVO sentences are more frequently used than VOS sentences, and because the verb does not need to be morphologically altered in SVO order, Kaqchikel is referred to as a language that is possibly shifting, or has already shifted, from VOS to SVO.

In summary, there are two theories regarding modern Kaqchikel's syntactically basic word order, one of which argues for VOS, and the other, for SVO. Indeed, reflecting this situation, *World Atlas of Language Structure* (Haspelmath et al. 2005) assumes that Kaqchikel's basic word order is VOS, while *Ethnologue* (Lewis 2009) classifies the language as SVO (see also Broady 1984).

Based on psycho- and neurolinguistic studies, it is known that relative to a language's syntactically basic word order, its other word orders (= derived word orders) have syntactic structures that are more complex when produced or comprehended by the brain (Mazuka, Ito, and Kondo 2002; Miyamoto and Takahashi 2002; Ueno and Kluender 2003; Kaiser and Trueswell 2004; Tamaoka et al. 2005; Caplan, Chen and Waters 2008; Grewe et al. 2007; Kinno et al. 2008; Kanduboda and Tamaoka 2012). Many recognized sentence processing theories predict that derived word orders with a filler-gap dependency have a higher processing load during sentence comprehension than the corresponding syntactically basic word order (Pritchett and Whitman 1995; Gibson 2000; Hawkins 2004).

Another possible factor that greatly influences sentence processing load is the usage frequency. While it is widely known that, at the word level of processing, frequently used words take less time to process, it has also been reported that there are cases where the frequency of words and sentence structures also affects the sentence processing load (e.g., Trueswell, Tanenhaus and Kello 1993). Therefore, speakers are more proficient in sentence structures and words that are used frequently in their language and are more likely to process these with speed and accuracy. Sentence processing theories based on these premises have also been proposed (e.g., Jurafsky 1996; Crocker and Branis 2000).

Based on the information from these preceding studies, we can assume that (54a) and (54b) hold true.

(54) a. Other things being equal, a word order that corresponds to a simple syntactic structure has a lower load in sentence processing compared to one that corresponds to a complex syntactic structure.

b. Other things being equal, a word order that is more frequently used has a lower load in sentence processing compared to one that is less frequently used.

Taking into account the sentence processing load of Kaqchikel based on these two premises, we can draw the following hypotheses based on whether the syntactically basic word order of Kaqchikel is SVO or VOS. First, if the current Kaqchikel syntactically basic word order is SVO, both (54a) and (54b) predict that SVO takes a lower processing load compared to VOS. On the other hand, if the current Kaqchikel syntactically basic word order is indeed VOS, then three different cases can be drawn, based on which factor has a larger impact on the sentence processing load –

namely, the syntactic factor of (54a) or the frequency factor of (54b). First, if the frequency factor has a larger influence on the sentence processing load compared to the syntactic factor, then it is predicted that SVO takes a lower processing load, even if the syntactically basic word order is VOS. Next, if frequency and syntax have a similar level of influence, the two factors should cancel each other's influences, and the sentence processing load of VOS and SVO should be similar. Lastly, if syntactic complexity has a greater influence on sentence processing than frequency does, then it is predicted that VOS has a lower processing load compared to SVO. These predictions are summarized in (55) (wherein A < B represents A taking a lower processing load than B, and A = B represents a similar level of processing load).

(55) Prediction of Kaqchikel sentence processing load:
   a. If the syntactically basic word order is SVO: SVO < VOS

   b. If the syntactically basic word order is VOS, and
      i. Frequency is more influential: SVO < VOS
      ii. Frequency and syntactic structure have a similar level of influence: SVO = VOS
      iii. Syntactic structure is more influential: SVO > VOS

In order to verify whether any of these predictions is correct, Koizumi et al. (2014) conducted a sentence processing experiment using Kaqchikel transitive verb sentences (see also Kiyama et al. 2013). Grammatically correct and semantically cohesive transitive sentences (56) were arranged in three different word orders (VOS, SVO, VSO).

(56) a. [VOS] *Xuchöy ri chäj ri ajanel.*
   cut DET pine.tree DET carpenter

   b. [VSO] *Xuchöy ri ajanel ri chäj.*
   cut DET carpenter DET pine.tree

   c. [SVO] *Ri ajanel xuchöy ri chäj.*
   DET carpenter cut DET pine.tree
   'The carpenter cut the pine tree.'

A sentence plausibility judgment task was administered. In this task, the stimulus sentences were phonologically presented in a random order to the participants through headsets. The participants were asked to judge whether each sentence was semantically correct and to push a "yes" button (correct sentence) or "no" button (incorrect sentence) as quickly and accurately as possible according to their judgment. The time from the beginning of each stimulus sentence until the button press was measured as the reaction time.

The results of the experiment showed that VOS (3403 ms) had a lower processing load than SVO (3559 ms) and VSO (3601 ms). This is in agreement with the

prediction in (55b-iii) that “VOS is the syntactically basic word order, and syntactic structure is more influential” and disagrees with the prediction of (55a) that “SVO is the syntactically basic word order.” Therefore, the results suggest that the syntactically basic word order of modern Kaqchikel is VOS, which is in alignment with the traditional analysis (García Matzar and Rodríguez Guaján 1997: 333). That is to say that if indeed a part of the modern Kaqchikel community is currently shifting from using VOS to SVO as the syntactically basic word order, this shift has not yet been reflected in the internal grammar of the majority of native Kaqchikel speakers.

# 4 Conclusion

In this chapter, we have reviewed a number of studies in the area of experimental syntax in the broad sense. They amply demonstrate that experimental techniques, if combined with insightful theoretical hypotheses, are quite useful in revealing the nature of the cognitive system of human language as well as that of related performance systems. In particular, various online indices obtained through experiments provide us with invaluable data in identifying possible and impossible representations as well as underlying constraints in free word order languages such as Japanese, in which surface constituent orders do not necessarily provide sufficient information to do so.

It is evident from the discussion above that in order to investigate the nature of the human parser (the first category of experimental syntax, defined above as (I)), researchers need a firm knowledge of individual grammars and linguistic theories. Conversely, it is also evident that experimental studies constitute an important testing ground for the evaluation of competing linguistic theories/hypotheses (the second category of experimental syntax, defined above as (II), and also referred to as experimental syntax in the narrow sense) on the condition that the nature of the human parser is reasonably well understood. This is because experimental data are interpreted under certain assumptions based on studies of the first category of experimental syntax. It is thus essential for researchers in the different disciplines to share their findings, to relate them to one another, and more ideally, to integrate the fields into a unified approach in order to elucidate how the brain enables language (Phillips 2004).

# Acknowledgments

I would like to thank Mineharu Nakayama for giving me the opportunity to contribute this chapter to this volume. Part of this work was supported by a Grant-in-Aid for Scientific Research (S) entitled “A field-based, cognitive neuroscientific study of the processing of OS-type languages” (No. 22222001, PI: Masatoshi Koizumi) from the Japan Society for the Promotion of Science.

# References


Aissen, Judith L. 1992. Topic and focus in Mayan. *Language* 68. 43–80.

Ajsivinac Sian, Juan Esteban, Lolmay Pedro Oscar García Mátzar, Martín Chacach Cutzal and Ixkusamil Estela Alonzo Guaján. 2004. *Gramática Descriptiva del Idioma Maya Kaqchikel: Rutzijoxik Rucholik Ri Kaqchikel Ch'ab'äl*. Chimaltenango, Guatemala: Comunidad Lingüística Kaqchikel de la Academia de las Lenguas Mayas de Guatemala.

Belletti, Adriana (ed.). 2004. *The cartography of syntactic structures. Vol. 3, Structures and beyond*. Oxford: Oxford University Press.

Bock, J. Kathryn and David E. Irwin. 1980. Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior* 19. 467–484.

Bock, J. Kathryn and Richard K. Warren. 1985. Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21. 47–67.

Broady, Jill. 1984. Some problems with the concept of basic word order. *Linguistics* 22. 711–736.

Brown, R. McKenna, Judith M. Maxwell and Walter E. Little. 2006. *¿La ütz awäch?: Introduction to Kaqchikel Maya language*. Austin: University of Texas Press.

Caplan, David, Evan Chen and Gloria Waters. 2008. Task-dependent and task-independent neurovascular responses to syntactic processing. *Cortex* 44. 257–275.

Chomsky, Noam. 1957. *Aspects of the theory of syntax*. Cambridge MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chujo, Kazumitsu. 1983. Nihongo tanbun-no rikai katei – Bunrikai sutoratejī no sōgo kankei [The interrelationships among strategies for sentence comprehension]. *Japanese Journal of Psychology* 54. 250–256.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A Cross-linguistic perspective*. Oxford: Oxford University Press.

Cinque, Guglielmo (ed.) 2002. *The cartography of syntactic structures*, *Vol. 1: Functional structure in DP and IP*. Oxford: Oxford University Press.

Cinque, Guglielmo and Luigi Rizzi. 2008. The cartography of syntactic structure. *Studies in Linguistics: CISCL Working Papers* 2. 43–59.

Clifton, Charles and Frazier, Lyn. 2004. Should given information come before new? Yes and no. *Memory and Cognition* 32. 886–895.

Coon, Jessica. 2010. VOS as predicate fronting in Chol Mayan. *Lingua* 120. 345–378.

Cowper, Elizabeth. 1999. Grammatical aspect in English. *The Linguistic Review* 16. 205–226.

Crocker, Matthew W. and Thorsten Brants. 2000. Wide-coverage probabilistic sentence processing. *Journal of Psycholinguistic Research* 29. 647–669.

De Vincenzi, Marica. 1991. Filler-gap dependencies in a null subject language: Referential and non-referential WHs. *Journal of Psycholinguistic Research* 20. 197–213.

Dubinsky, Stanley and Shoko Hamano. 2003. Case checking by AspP: The syntax and semantics of predicative postpositions. In William, McClure (ed.), *Japanese/Korean Linguistics* 12, 231–242. Stanford: CSLI Publications.

England, Nora C. 1991. Changes in basic word order in Mayan languages. *International Journal of American Linguistics* 57. 446–486.

Erdocia, Kepa, Itziar Laka, Anna Mestres-Missé and Antoni Rodriguez-Fornells. 2009. Syntactic complexity and ambiguity resolution in a free word order language: Behavioral and electrophysiological evidences from Basque. *Brain and Language* 109. 1–17.

Ernst, Thomas. 2002. *The syntax of adjuncts*. Cambridge: Cambridge University Press.

Ferreira, Victor S. and Hiromi Yoshita. 2003. Given-new ordering effect on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research* 32. 669–692.

Frazier, Lyn and Giovanni B. Flores d'Arcais. 1989. Filler-driven parsing: A study of gap-filling in Dutch. *Journal of Memory and Language* 28. 331–344.

Fukui, Naoki. 1995. *Theory of projection in syntax.* Stanford, CA: CSLI Publications.

García Matzar, Lolmay Pedro and Pakal B'alam José Obispo Rodríguez Guaján. 1997. *Rukemik ri Kaqchikel chi': Gramática kaqchikel*. Guatemala, Guatemala: Cholsamaj.

Gibson, Edward. 2000. Dependency locality theory: A distance-based theory of linguistic complexity. In Alec Marantz, Yasushi Miyashita and Wayne O'Neil (eds.), *Image, language, brain: Papers from the first Mind Articulation Project Symposium*, 95–126. Cambridge, MA: MIT Press.

Grewe, Tanja, Ina Bornkessel-Schlesewsky, Stefan Zysset, Richard Wiese, D. Yves von Cramon and Matthias Schlesewsky. 2007. The role of the posterior superior temporal sulcus in the processing of unmarked transitivity. *NeuroImage* 35. 343–352.

Hagiwara, Hiroko, Takahiro Soshi, Masami Ishihara and Kuniyasu Imanaka. 2007. A topographical study on the event-related potential correlates of scrambled word order in Japanese complex sentences. *Journal of Cognitive Neuroscience* 19. 175–193.

Haspelmath, Martin, Matthew S. Dryer, David Gil and Bernard Comrie (eds.) 2005. *The world atlas of language structures*. Oxford: Oxford University Press.

Haviland, Susan E. and Herbert H. Clark. 1974. What's new? Acquiring information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior* 13. 512–521.

Hawkins, John A. 2004. *Efficiency and complexity in grammars.* Oxford: Oxford University Press.

Hoji, Hajime. 1985. *Logical form constraints and configurational structures in Japanese*. Seattle, WA: University of Washington dissertation.

Imamura, Satoshi and Masatoshi Koizumi. 2008a. Bunshori-ni okeru jōhō kōzō-to tōgo kōzō-no kōgosayō-ni tsuite [On the interaction between information structure and syntactic structure in sentence processing]. *Proceedings of the 136th Conference of the Linguistics Society of Japan*, 212–217. Kyoto: Linguistics Society of Japan.

Imamura, Satoshi and Masatoshi Koizumi. 2008b. Bunrikai-ni okeru jōhō kōzō-to tōgo kōzō-no kōgosayō-ga sōjiru taimingu-ni tsuite [On the time course of the interaction between information structure and syntactic structure in sentence comprehension]. *Proceedings of the 137th Conference of the Linguistics Society of Japan*, 92–97. Kyoto: Linguistics Society of Japan.

Imamura, Satoshi and Masatoshi Koizumi. 2011. A centering analysis of word order in Japanese. *Tohoku Studies in Linguistics* 20. 59–74.

Inubushi, Tomoo, Kazuki Iijima, Masatoshi Koizumi and Kuniyoshi L. Sakai. 2012. Left inferior frontal activations depending on the canonicity determined by the argument structures of ditransitive sentences: An MEG study. *PloS ONE* 7(5). e37192.

Inui, Toshio, Yukio Otsu, Shigeki Tanaka, Tomohisa Okada, Sadahiko Nishizawa and Junji Konishi. 1998. A functional MRI analysis of comprehension processes of Japanese sentences. *Neuroreport* 9. 3325–3328.

Ishida, Jun. 1999. *Bun-no yomiyasu-sa-to hyōgen keishiki-to-no kankei – gojun, tōgo kōzō, oyobi daiyōkei siyō-ni kansuru kentō* [The relation between the readability of sentences and the form of expressions relating to word order, syntactic structure, and the use of pronouns]. Kobe: Kobe University of Commerce.

Jurafsky, Daniel. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20. 137–194.

Kaiser, Elsi and John C. Trueswell. 2004. The role of discourse context in the processing of a flexible word-order language. *Cognition* 94. 113–147.

Kamide, Yuki, Gerry T. M. Altmann and Sarah L. Haywood. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language* 49. 133–156.

Kanduboda, Arachchige Buddhika Prabath and Katsuo Tamaoka. 2012. Priority information determining the canonical word order of written Sinhalese sentences. *Open Journal of Modern Linguistics* 2(1). 26–33.

Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8. 63–99.

Kim, Jungho. 2012. Kankokugo kakimazegojunbun-no puraimingu kōka [Priming effects in scrambled sentences in Korean]. *Culture* 75. 141–156.

Kim, Jungho, Masatoshi Koizumi, Naho Ikuta, Yuichiro Fukumitsu, Naoki Kimura, Kazuki Iwata, Jobu Watanabe, Satoru Yokoyama, Shigeru Sato, Kaoru Horie and Ryuta Kawashima. 2009. Scrambling effects on the processing of Japanese sentences: An fMRI study. *Journal of Neurolinguistics* 22(1). 151–166.

Kimura, Naoki. 2004. *An analysis of Japanese multiple adverbial constructions: Configurationality of modality, negation, and VP.* Tokyo: International Christian University Master's thesis.

Kimura, Naoki, Jungho Kim, and Masatoshi Koizumi. 2005. Sentence processing and phrase structural determinacy of aspect in Japanese. *Lexicon Forum* 1. 133–161.

Kinno, Ryuta, Mitsuru Kawamura, Seiji Shioda and Kuniyoshi L. Sakai. 2008. Neural correlates of non-canonical syntactic processing revealed by a picture-sentence matching task. *Human Brain Mapping* 29. 1015–1027.

Kitagawa, Yoshihisa. 1986. *Subjects in Japanese and English*. Amherst, MA: University of Massachusetts dissertation.

Kiyama, Sachiko, Katsuo Tamaoka, Jungho Kim and Masatoshi Koizumi. 2013. Effect of animacy on word order processing in Kaqchikel Maya. *Open Journal of Modern Linguistics* 3(3).

Koizumi, Masatoshi. 1993. Modal phrase and adjuncts. In Patricia M. Clancy (ed.), *Japanese/Korean Linguistics* 2, 409–428. Stanford: CSLI Publications.

Koizumi, Masatoshi, Jungho Kim, Naoki Kimura, Satoru Yokoyama, Shigeru Sato, Kaoru Horie and Ryuta Kawashima. 2012. Left inferior frontal activations differentially modulated by scrambling in ditransitive sentences. *The Open Medical Imaging Journal* 6. 70–79.

Koizumi, Masatoshi and Katsuo Tamaoka. 2004. Cognitive processing of Japanese sentences with ditransitive verbs. *Gengo Kenkyu* 125. 173–190.

Koizumi, Masatoshi and Katsuo Tamaoka. 2006. Bunkaisekijikken-niyoru nihongo-fukushirui-no gojunno hantei [Determination of basic word order of adverbs in Japanese by a sentence-processing experiment]. *Cognitive Studies* 13. 392–403.

Koizumi, Masatoshi and Katsuo Tamaoka. 2010. Psycholinguistic evidence for the VP-internal subject position in Japanese. *Linguistic Inquiry* 41. 663–680.

Koizumi, Masatoshi, Yoshiho Yasugi, Katsuo Tamaoka, Sachiko Kiyama, Jungho Kim, Lolmay Pedro Oscar García Matzar and Juan Esteban Ajsivinac Sian. 2014. On the (non-)universality of the preference for subject-object word order in sentence comprehension: A sentence processing study in Kaqchikel Maya. *Language* 90. 722–736.

Koopman, Hilda and Dominique Sportiche. 1988. The positions of subjects. *Lingua* 85. 211–258.

Kubo, Takuya, Hajime Ono, Mikihiro Tanaka, Masatoshi Koizumi and Hiromu Sakai. 2012. How does animacy affect word order in a VOS language. Poster presented at *the 25th Annual CUNY Conference on Human Sentence Processing*, New York, March 16, 2012.

Kuno, Susumu. 1978. *Danwa-no bunpō* [Grammar of discourse]. Tokyo: Taishuukan.

Kuroda, S.-Y. 1988. Whether we agree or not. *Lingvisticae Investigationes* 12. 1–47.

Kutas, Marta and Steven A. Hillyard. 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207. 203–204.

Lewis, M. Paul (ed.) 2009. *Ethnologue: Languages of the world, Sixteenth edition*. Dallas, TX: SIL International. Online version: http://www.ethnologue.com/.

Marantz, Alec. 1992. Case and licensing. In German Westphal, Benjamin Ao and Hee-Rahk Chae (eds.), *Proceedings of the Eighth Eastern States Conference on Linguistics (ESCOL '91)*, 234–253. Ithaca, NY: Cornell University, Cornell Linguistics Club Publications. Reprinted in Eric J. Reuland (ed.), *Arguments and case: Explaining Burzio's generalization*, 11–30, Amsterdam: John Benjamins, 2000.

Marantz, Alec. 2005. Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review* 22. 429–445.

Marslen-Wilson, William D. 1975. Sentence perception as an interactive parallel process. *Science* 189. 226–228.

Mazuka, Reiko, Kenji Itoh and Tadahisa Kondo. 2002. Costs of scrambling in Japanese sentence processing. In Mineharu Nakayama (ed.), *Sentence processing in East Asian languages*, 131–166. Stanford, CA: CSLI Publications.

McClure, William. 1994. *Syntactic projections of the semantics of aspect*. Ithaca, NY: Cornell University dissertation.

Minami, Fujio. 1974. *Gendai nihongo-no kōzō* [Structure of contemporary Japanese]. Tokyo: Taishukan.

Miyagawa, Shigeru. 1989. *Structure and case marking in Japanese*. San Diego: Academic Press.

Miyagawa, Shigeru. 2001. The EPP, scrambling, and Wh-in-situ. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 293–338. Cambridge, MA: MIT Press.

Miyagawa, Shigeru. 2003. A-movement scrambling and options without optionality. In Simin Karimi (ed.), *Word order and scrambling*, 177–200. Malden, MA: Blackwell.

Miyagawa, Shigeru. 2006. Locality in syntax and floated numeral quantifiers in Japanese and Korean. In Timothy J. Vance and Kimberly Jones (eds.), *Japanese/Korean linguistics 14*, 270–282. Stanford, CA: CSLI Publications.

Miyagawa, Shigeru and Koji Arikawa. 2007. Locality in syntax and floating numeral quantifiers. *Linguistic Inquiry* 38. 645–670.

Miyamoto, Edson T. and Shoichi Takahashi. 2002. Sources of difficulty in the processing of scrambling in Japanese. In Mineharu Nakayama (ed.), *Sentence processing in East Asian languages*, 167–188. Stanford, CA: CSLI Publications.

Mulders, Iris Corine Margarethe Christine. 2002. *Transparent parsing: Head-driven processing of verb-final structures*. Utrecht: Landelijke Onderzoekschool Taalwetenschap dissertation.

Nakano, Yoko, Claudia Felser and Herald Clahsen. 2002. Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research* 31. 531–571.

Nakau, Minoru. 1980. Bunfukushi-no hikaku [The comparison of sentential adverbs]. In Tetsuya Kunihiro (ed.), *Nichi-eigo hikaku kōza 2: Bunpō* [Lectures on Japanese-English comparative studies 2: Grammar], 157–219. Tokyo: Taishukan.

Nakayama, Mineharu and Richard Lewis. 2001. Similarity interference and scrambling in Japanese. *Injiguahak Jakop* [Journal of Cognitive Science] 1. 39–53.

Nakayama, Mineharu (ed.) 2002. *Sentence processing in East Asian languages*. Stanford, CA: CSLI Publications.

Nemoto, Naoko. 1999. Scrambling. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 121–153. New York: Blackwell.

Noda, Hisashi. 1984. Fukushi-no gojun [Word order of adverbs]. *Nihongo Kyōiku* 52. 79–90.

Phillips, Colin. 2004. Linguistics and linking problems. In Mabel L. Rice and Steven F. Warren (eds.), *Developmental language disorders: From phenotypes to etiologies*, 241–287. Mahwah, NJ: Lawrence Erlbaum Associates.

Preminger, Omer. 2011. *Agreement as a fallible operation*. Cambridge, MA: MIT dissertation.

Primus, Beatrice. 1999. *Cases and thematic roles*. Tübingen: Niemeyer.

Pritchett, Bradley L. 1988. Garden path phenomena and the grammatical basis of language processing. *Language* 64. 539–576.

Pritchett, Bradley L. 1991. Head position and parsing ambiguity. *Journal of Psycholinguistic Research* 20. 251–270.

Pritchett, Bradley L. 1992. *Grammatical competence and parsing performance*. Chicago: The University of Chicago Press.

Pritchett, Bradley L. and John B. Whitman. 1995. Syntactic representation and interpretive preference. In Reiko Mazuka and Noriko Nagai (eds.), *Japanese sentence processing*, 65–76. Hillsdale, NJ: Lawrence Erlbaum.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.) *Elements of grammar*, 281–337. Amsterdam: Kluwer.

Rizzi, Luigi (ed.) 2004. *The cartography of syntactic structures*. Vol. 2, *The structure of CP and IP*. Oxford: Oxford University Press.

Rösler, Frank, Thomas Pechmann, Judith Streb, Brigitte Roeder and Erwin Hennighausen. 1998. Parsing of sentences in a language with varying word order: Word-by-word variations of processing demands are revealed by event-related brain potentials. *Journal of Memory and Language* 38. 150–176.

Saito, Mamoru. 1985. *Some asymmetries in Japanese and their theoretical implications*. Cambridge, MA: MIT dissertation.

Saito, Mamoru and Hajime Hoji. 1983. Weak crossover and move α in Japanese. *Natural Language and Linguistic Theory* 1. 245–259.

Sekerina, Irina A. 1997. The syntax and processing of scrambling constructions in Russian. New York: City University of New York dissertation.

Shibatani, Masayoshi. 1977. Grammatical relations and surface cases. *Language* 53. 789–809.

Shibatani, Masayoshi. 1978. Mikami Akira and the notion of 'subject' in Japanese grammar. In John Hinds and Irwin Howards (eds.), *Problems in Japanese syntax and semantics*. Tokyo: Kaitakusha.

Stromswold, Karin, David Caplan, Nathaniel Alpert and Scott Rauch. 1996. Localization of syntactic comprehension by Positron Emission Tomography. *Brain and Language* 52. 452–473.

Tada, Hiroaki. 1993. A/A-bar partition in derivation. Cambridge, MA: MIT dissertation.

Takubo, Yukinori. 1987. Tōgokōzō-to bunmyaku jōhō [Syntactic structure and contextual information]. *Nihongogaku* 6(5). 37–48.

Tamaoka, Katsuo, Hiromu Sakai, Jun-ichiro Kawahara, Yayoi Miyaoka, Hyunjung Lim and Masatoshi Koizumi. 2005. Priority information used for the processing of Japanese sentences: Thematic roles, case particles or grammatical functions? *Journal of Psycholinguistic Research* 34. 273–324.

Thráinsson, Höskuldur. 1996. On the (non)-universality of functional categories. In Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson and Jan-Wouter. Zwart (eds.), *Minimal ideas: Syntactic studies in the minimalist framework*, 253–281. Amsterdam: John Benjamins.

Tichoc Cumes, Rosalino, Juan Esteban Ajsivinac Sian, Lolmay Pedro Oscar García, Ixchel Carmelina Espantzay, Martín Chacach Cutzal and Estela Alosno Guajan. 2000. *Runuk'ul pa Rub'eyal Rutz'ib'axik ri Kaqchikel Ch'ab' äl: Grammática Normativa del Idioma Maya Kaqchikel*. Chimaltenango, Guatemala: Comunidad Lingüística Kaqchikel de la Academia de las Lenguas Mayas de Guatemala.

Travis, Lisa Demena. 2005. Articulated vPs and the computation of aspectual classes. In Paula Kempchinsky and Roumyana Slabakova (eds.), *Aspectual Inquiries*, 69–93. Dordrecht: Springer.

Trueswell, John C., Michael K. Tanenhaus and Christopher Kello. 1993. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition* 19. 528–553.

Ueno, Mieko and Robert Kluender. 2003. Event-related brain indices of Japanese scrambling. *Brain and Language* 86. 243–271.

Vande Kopple, William J. 1982. The given-new strategy of comprehension and some natural expository paragraphs. *Journal of Psycholinguistic Research* 11. 501–520.

Weyerts, Helga, Martina Penke, Thomas F. Münte, Hans-Jochen Heinze and Harald Clahsen. 2002. Word order in sentence processing: An experimental study of verb placement in German. *Journal of Psycholinguistic Research* 31. 211–268.

Barış Kahraman and Hiromu Sakai

# 14 Relative clause processing in Japanese: Psycholinguistic investigation into typological differences

## 1 Introduction

Relative clauses have been one of the most widely studied constructions since the pioneering study by Keenan and Comrie (1977), in the field of language typology. In the field of psycholinguistics, too, the processing of relative clauses has been one of the most extensively studied topics in the past decades. It is therefore natural for psycholinguists to bring typological perspectives into research on relative clause processing. That is, those who are investigating the processing of typologically distinct languages such as English and Japanese are interested in the influence of typological factors on language processing. Researchers initially focused on surface syntactic factors such as constituent order or structural complexities and then gradually shifted their attentions to various other factors, including semantic factors such as animacy of noun phrases or pragmatic factors such as the discourse functions of relative clauses. Importantly, this shift of attention was in line with development of theories on relative clause processing. Some theories attribute sources of processing difficulties to the process of *filler-gap dependency formation*. Other theories assume that *learning through experience*, namely *statistical learning*, is an essential aspect of human sentence processing that determines processing difficulties. In this chapter, we briefly review the history of Japanese relative clause processing research and discuss what we have found and what is left for future research. We discuss why it is important to consider various types of typological factors.

In quite a few numbers of languages in the world, it has been well-documented that there exist an asymmetry in the comprehension of subject relative clauses (SRCs) like (1a) and (2a) and object relative clauses (ORCs) like (1b) and (2b).

(1) English:
a. SRC: *The student* [*who saw the teacher*] *was waiting for the bus.*
b. ORC: *The student* [*who the teacher saw*] *was waiting for the bus.*

(2) Japanese:

a. SRC:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| [*Sensei* | *o* | *mita*] | *gakusei* | *wa* | *basu* | *o* | *matteita.* |
| teacher | ACC | saw | student | TOP | bus | ACC | was.waiting |

‘The student who saw the teacher was waiting for the bus.’

b. ORC:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| [*Sensei* | *ga* | *mita*] | *gakusei* | *wa* | *basu* | *o* | *matteita.* |
| teacher | NOM | saw | student | TOP | bus | ACC | was.waiting |

‘The student who the teacher saw was waiting for the bus.’

SRCs are generally easier to process than ORCs (e.g., Chinese: Lin 2006; Dutch: Mak, Vonk and Schfriers 2006; English: Staub 2010; French: Cohen and Mehler 1996; German: Schriefers, Friederici and Kühn 1995; Hungarian: MacWhinney and Pleh 1988; Japanese: Ueno and Garnsey 2008; Korean: Kwon et al. 2013; Spanish: Betancort, Carreiras and Sturt 2009; Turkish: Kahraman et al. 2010).[1]

It should be emphasized that relative clauses of languages in the world show both typological similarities and differences. For example, even though relative clauses are used to 'modify' the content of head-noun in a broad sense in all of the languages mentioned above, they fulfill such function in quite different ways. In a head-initial language like English, the head-noun (*the teacher*) comes before the relative clause as in (1). In a head-final language like Japanese, on the other hand, the relative clause comes before the head-noun (*gakusei*) as in (2). Such cross-linguistic differences are important not only from typological perspective, but also from the psycholinguistic perspective. Moreover, a relative pronoun (e.g., *who*) unambiguously signals the existence of a relative clause in English. Japanese, however, does not have a relative pronoun or a relative clause marker. Therefore, Japanese relative clauses are temporarily ambiguous between relative clauses and other types of subordinate/matrix clauses until the head-noun is encountered. English speaking readers/listeners therefore understand that they are reading/hearing a relative clause when they encounter a relative pronoun, whereas Japanese speaking readers/listeners cannot realize the existence of a relative clause until they encounter a head-noun. The exploration of the source of processing asymmetry between SRCs and ORCs in typologically different languages, thus provides important insights into characterization of human cognitive mechanisms for language processing by revealing their universal and language-specific aspects.

In order to explain the processing asymmetry between SRCs and ORCs, researchers have proposed various accounts (see also Sakamoto, this volume). The recent accounts can roughly be classified into two groups according to the theories of sentence processing they are based on. One is based on the theories of filler-gap dependency formation. They mainly examine the influence on relative clause processing by surface syntactic factors, such as constituent order or structural complexity.[2] The other is based on

**1** Actually, a few studies reported ORC advantage over SRC in languages like Chinese (Hsiao and Gibson 2003; Lin 2014) and Basque (Carreiras et al. 2010), but in general, SRCs are easier than ORCs.
**2** There is also another complexity based account which is called *Similarity Interference Hypothesis*. This account basically assumes that the processing difficulty of ORCs stems from the repetition of same types of noun phrases in English (Gordon, Hendrick and Johnson 2001). In SRCs, the thematic roles can be easily assigned because no extra noun phrase intervenes between the relative clause verb and the head-noun. In ORCs, on the other hand, two noun phrases should be held in the memory until the thematic roles are assigned. In Japanese, thematic roles are clearly marked by case markers and semantic ambiguities do not arise except in limited cases. Therefore, we will not discuss the Similarity Interference Hypothesis in this chapter. See also Nakayama, Vasishth and Lewis (2006) on similarity interference in Japanese.

various kinds of factors learned through experience, such as; predictability of upcoming elements, frequency of structures, animacy of noun phrases, and discourse functions of relative clauses. We will discuss these accounts in detail. However, since the great majority of the accounts have been proposed based on the data from relative clause processing in English and other European languages, it is quite difficult to evaluate the validity of these accounts by just looking at data from European languages. In this context, cross-linguistic investigation of relative clause processing, especially in languages with typological differences such as Japanese, gains more importance because it allows us to distinguish the validity and the universality of these accounts, and construct a more accurate model of human sentence processing mechanisms.

In what follows, we will first discuss accounts based on the theories on filler-gap dependency formation, and specifically test the validity of Dependency Locality Theory (Gibson 1998, 2000) and Structural Distance Hypothesis (O'Grady 1997) in Japanese. Then, in section 3, we will review accounts based on other factors and the theories on statistical learning through experience in Japanese. Finally, in section 4, we will discuss the limitations and possibilities of future directions of relative clause processing studies in Japanese.

# 2 Relative clause processing and filler-gap dependency

First we will examine accounts for relative clause processing based on two representative theories on filler-gap dependency formation: the Dependency Locality Theory and the Structural Distance Hypothesis. We will show that both of these theories make the same prediction for the processing difficulty of relative clauses in English, whereas these theories make a different prediction in Japanese. We will then briefly introduce studies that attempt to verify these predictions and that examine the influence of syntactic factors on relative clause processing in Japanese. Our focus is on Japanese native speakers' relative clause processing in this chapter, but those who are interested in the evaluation of the two accounts above in L2 Japanese processing should be referred to Sawasaki and Kashiwagi-Wood's chapter in this volume.

## 2.1 Dependency Locality Theory and Structural Distance Hypothesis

The concept of *filler-gap dependency* plays an important role in the sentence processing literature (see also Sakamoto's chapter in this volume). An argument of a verb can be displaced from its original position, appearing in another position in the

sentence, e.g., WH-questions in English. In the field of sentence processing, a displaced element is called a *filler*, and in its original position is a *gap*. A filler-gap dependency is found in many different constructions, such as clefts, relative clauses, topic sentences, WH-questions, and so on. In order to understand the meaning of these kinds of sentences, listeners/readers set up an association between the filler and the gap. This refers to the filler-gap dependency (e.g., Fodor 1989). For instance, *student* is the filler, and the underline shows the gap position in English examples below.[3]

(3) a. SRC: *The student$_i$ who* ___$_i$ *saw the teacher…*
 b. ORC: *The student$_i$ who the teacher saw* ___$_i$ …

According to the Dependency Locality Theory, the number of discourse referents (words) between the filler and the gap is the main source of the processing difficulty of relative clauses (Gibson 1998, 2000). This type of account is also conventionally called Linear Distance Hypothesis (Ueno and Garnsey 2008). Gibson argues that the listeners/readers need to hold the filler in working memory and retrieve it at the gap position. SRCs are easier to comprehend than ORCs because the number of discourse referents intervening between the filler and the gap are fewer in SRCs than ORCs (see Sawasaki and Kashiwagi-Wood's chapter). Since the working memory load is heavier for ORCs, the filler is harder to retrieve in ORCs compared to SRCs at the relative clause verb *saw* in (3).

On the other hand, the Structural Distance Hypothesis (O'Grady 1997; Hawkins 1999, 2004) considers that the number of the syntactic nodes between the filler and the gap, or the embedding depth of the gap, determines the processing load. According to O'Grady (1997: 136), structural complexity increases by the number of syntactic nodes between the filler and the gap, and this causes a processing disadvantage.

As illustrated in Figure 1, there are three nodes between the filler and the gap in SRCs, whereas there are four nodes in ORCs. Since the computational complexity is heavier in ORCs compared to SRCs, this hypothesis predicts that SRCs are easier to comprehend than ORCs.

Both accounts predict that SRCs should be easier to process than ORCs, and they are indistinguishable in the head-initial languages like English. However, in a head-final language like Japanese, these two theories make different predictions for the processing difficulty of SRCs and ORCs. The Dependency Locality Theory predicts that ORCs should be easier to process than SRCs, because the memory load is heavier for SRCs, i.e., more intervening words between the filler and the gap in SRCs than ORCs.

3 Some researchers accept a relative pronoun as the filler in English. In many languages, no relative pronoun is used but fillers do exist. In this chapter, we assume the head-noun as the filler.

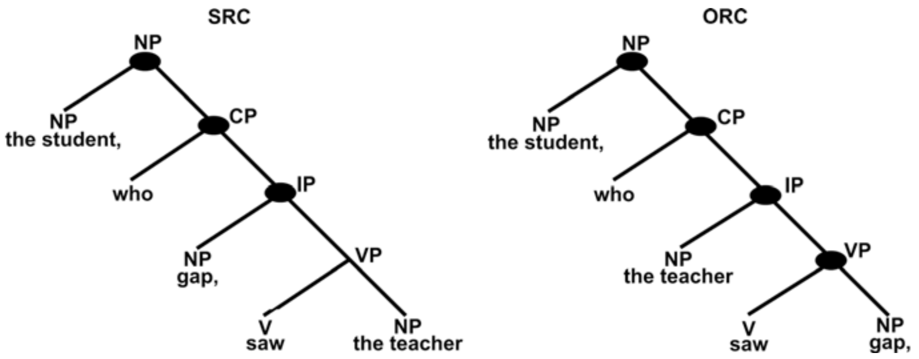



**Figure 1:** Structural distance of subject and object relative clauses in English

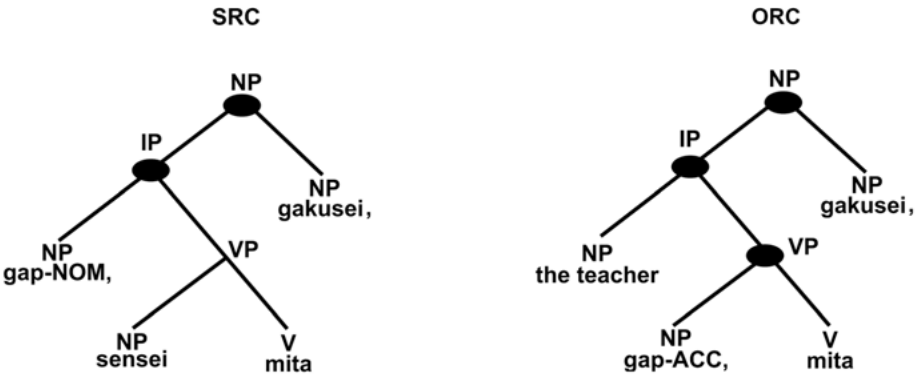



**Figure 2:** Structural distance of subject and object relative clauses in Japanese[4]

(4) a. SRC: [___$_i$ *Sensei* *o* *mita*] *gakusei$_i$* *wa*...
teacher ACC saw student TOP
'The student who saw the teacher...'

b. ORC: [*Sensei* *ga* ___$_i$ *mita*] *gakusei$_i$* *wa*...
teacher NOM saw student TOP
'The student who the teacher saw...'

On the other hand, the number of syntactic nodes between the filler and the gap is fewer in SRCs than ORCs as illustrated in Figure 2. Therefore, unlike the Dependency Locality Theory, the Structural Distance Hypothesis predicts that SRCs should be easier to process than ORCs because SRCs are less structurally complex.

**4** Some researchers add a CP node over the IP node (e.g. Ishizuka 2005). In this chapter, it is not a crucial issue. We, thus, based on Murasugi (2000), assume that SRCs and ORCs are IP, and did not add a CP node.

In sum, these two theories make different predictions in Japanese, and this allows researchers to distinguish between competing hypotheses, which was impossible in English.

## 2.2 Relative clause processing in Japanese from the filler-gap dependency formation perspective

The theories of filler-gap dependency formation make different predictions about Japanese relative clause processing, and researchers previously examined the validity of these theories (e.g., Miyamoto and Nakamura 2003; Nakamura 2003; Ishizuka 2005; Ueno and Garnsey 2008; Sakamoto and Yasunaga 2009; Mitsugi, MacWhinney and Shirai 2010; Kahraman 2012). For example, Kahraman (2012) used test sentences as shown in (5), and compared the processing difficulty of SRCs and ORCs through a self-paced reading task. He compared the mean reading times of each word in the SRC and ORC conditions.

(5) a. SRC:

| | | | | | | |
|---|---|---|---|---|---|---|
| *Depaato* | *de* | *ryoosin* | *o* | *sagasiteita* | *kodomo* | *wa* |
| department store | LOC | parents | ACC | looking.for | child | TOP |

| | |
|---|---|
| *kyuuni* | *nakidasita.* |
| suddenly | cried |

'At the department store, the child who was looking for his parents suddenly began crying.'

b. ORC:

| | | | | | | |
|---|---|---|---|---|---|---|
| *Depaato* | *de* | *ryoosin* | *ga* | *sagasiteita* | *kodomo* | *wa* |
| department store | LOC | parents | NOM | looking.for | child | TOP |

| | |
|---|---|
| *kyuuni* | *nakidasita.* |
| suddenly | cried |

'At the department store, the child who the parents were looking for started crying.'

In both conditions, the test sentences started with a locative adverb. In the SRC condition, an accusative marked noun followed the adverb, and in the ORC condition, a nominative marked noun followed the adverb. The rest of the sentences were identical in the two conditions, in which a relative clause verb, a head-noun, an adverb and a matrix verb appeared in sequence. If the linear distance was a more important factor, the ORCs should be easier to process than SRCs at the head noun position, whereas the SRCs should be easier to process than the ORCs if the structural distance is a more important factor.

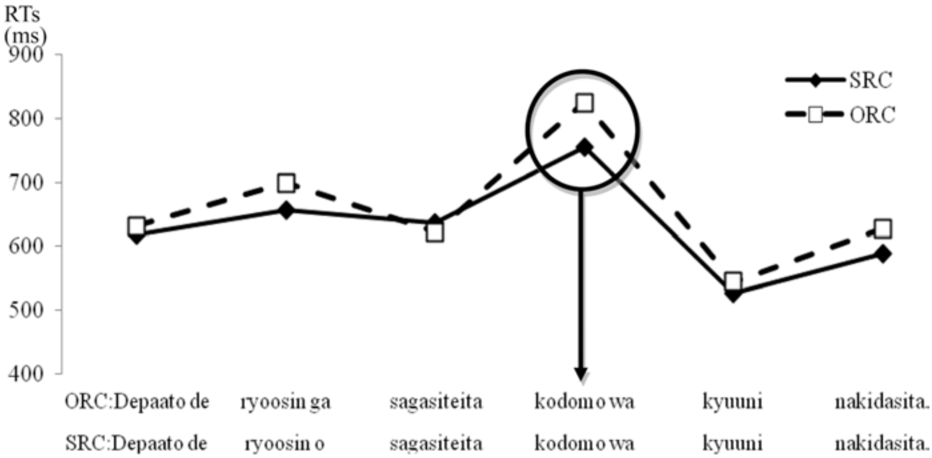



**Figure 3:** Reading times of SRCs and ORCs in Kahraman (2012) (The circle and the arrow show where the significant processing asymmetry was observed.)[5]

The results of the self-paced reading experiment showed that the reading times of the SRC's head-noun (*kodomo-wa*) were faster than that of the ORC, as shown in Figure 3.

Kahraman concluded that SRCs are easier to process than ORCs in Japanese, and the Structural Distance Hypothesis can account for the processing difficulty of ORCs (O'Grady 1997, Hawkins 1999, 2004), whereas the Dependency Locality Theory cannot (Gibson 1998, 2000).

In Japanese, various studies on relative clauses using different materials consistently reported that SRCs are easier to process than ORCs at the head-noun region (Ishizuka 2005; Kahraman et al. 2011a; Mitsugi, MacWhinney and Shirai 2010; Miyamoto and Nakamura 2003; Nakamura 2003; Sakamoto and Yasunaga 2009; Ueno and Garnsey 2008). For example, Miyamoto and Nakamura (2003) inserted an adverbial phrase between the relative clause verb and the sentence initial noun, and they manipulated the case marking of the head-noun of relative clause. The result of self-paced reading experiment showed that the head-noun of SRCs was read faster than that of ORCs, irrespective of the case marking of the head-noun. Based on these result, Miyamoto and Nakamura argued that the results cannot be explained by the Dependency Locality Theory. Ueno and Garnsey (2008) compared the processing difficulty of SRCs and ORCs through an experiment with a self-paced reading task and one with event related brain potentials (ERP). In both experiments, Ueno and Garnsey reported that the SRC were easier to process than the ORC at the relative

5 In the gloss, case markers are separately listed as independent morphemes, but in the experiments, case markers were presented with content words since they form phrases called *bunsetsu*. Therefore, we presented a content word and a case marker within the same region in the figures.

clause head noun position and they pointed out that the Structural Distance Hypothesis could successfully capture the processing asymmetry between the SRC and the ORC in Japanese. Unlike the above mentioned studies, Sakamoto and Yasunaga (2009) conducted a self-paced reading experiment using relative clause verbs which took dative case marked nouns as their arguments, such as *hanasu* 'talk', *au* 'met', *menkaisuru* 'met', etc. as in (6).

(6) a. SRC:

| | | | | | | |
|---|---|---|---|---|---|---|
| *Kisya* | *ni* | *menkaisita* | *giin* | *wa* | *kenryoku* | *ga* |
| reporter | DAT | met | senator | TOP | power | NOM |

| | | | |
|---|---|---|---|
| *subete* | *da* | *to* | *omotteita.* |
| everything | COP | that | was.thinking |

'The senator who met the reporter was thinking that the power is everything .'

b. ORC:

| | | | | | | |
|---|---|---|---|---|---|---|
| *Kisya* | *ga* | *menkaisita* | *giin* | *wa* | *kenryoku* | *ga* |
| reporter | DAT | met | senator | TOP | power | NOM |

| | | | |
|---|---|---|---|
| *subete* | *da* | *to* | *omotteita.* |
| everything | COP | that | was.thinking |

'The senator who the reporter met was thinking that the power is everything.'

Sakamoto and Yasunaga (2009) also found that SRCs were easier to process than ORCs, and concluded that the Structural Distance Hypothesis accurately captured the processing difficulty of relative clauses in Japanese. In addition to these studies, Kahraman (2012) and Mitsugi, MacWhinney and Shirai (2010) also compared the processing of SRCs and ORC by L2 learners of Japanese through self-paced reading experiments. The results of these studies showed that not only the native speakers but also Korean and Turkish speaking L2 learners of Japanese read SRCs faster than ORCs (see Sawasaki and Kashiwagi-Wood's chapter for detailed discussion of L2 relative clause processing in Japanese). Overall, these studies show that SRCs are generally easier to process than ORCs in Japanese. It means that the Structural Distance Hypothesis makes a valid prediction for Japanese relative clause processing but the Dependency Locality Theory cannot.

So far, we have shown that the studies on Japanese relative clause processing made a very important contribution to the field of sentence processing. However, these studies have mainly examined surface syntactic factors such as linear distance or structural distance and left many other potentially important factors unexamined. Accounts based on the theories of filler-gap dependency formation are valid only if other factors that potentially influence the processing load were controlled. Notice, however, that SRCs and ORCs which were used as experimental material in these

studies, repeated below, contain quite a few differences other than linear or structural distance between the filler and the gap.

(5) a. SRC:

| | | | | | | |
|---|---|---|---|---|---|---|
| *Depaato* | *de* | *ryoosin* | *o* | *sagasiteita* | *kodomo* | *wa* |
| department store | LOC | parents | ACC | looking.for | child | TOP |

| | |
|---|---|
| *kyuuni* | *nakidasita.* |
| suddenly | cried |

'At the department store, the child who was looking for his parents suddenly began crying.'

b. ORC:

| | | | | | | |
|---|---|---|---|---|---|---|
| *Depaato* | *de* | *ryoosin* | *ga* | *sagasiteita* | *kodomo* | *wa* |
| department store | LOC | parents | NOM | looking.for | child | TOP |

| | |
|---|---|
| *kyuuni* | *nakidasita.* |
| suddenly | cried |

'At the department store, the child who the parents were looking for started crying.'

Excluding the sentence initial adverbs or adjectives, the SRC and the ORC start from noun phrases with different case markers, a noun phrase with an accusative case marker in SRC and a noun phrase with a nominative case marker in ORC. It is well know that case markers play a very important role not only in determining grammaticality or semantic interpretation but also in processing Japanese sentences. Previous studies on Japanese sentence processing showed that readers use the case marker information very quickly and effectively to make predictions about the argument structure even before the verb appears (e.g., Kamide, Altmann and Haywood 2003; Miyamoto 2002; Yamashita 1997). It is conceivable that these case markers would elicit different predictions about the upcoming elements and/or structure. This kind of case marking difference and prediction mechanisms should be taken into consideration in Japanese relative clause processing studies. See Sakamoto's chapter on expectancy driven processing.

Second, previous studies have shown that the frequency of occurrence of SRCs and ORCs in corpora correlates with ease of relative clause processing in English (Reali and Christiansen 2007). The frequency of SRCs and ORCs might also be different and might have caused the processing asymmetry in Japanese. This possibility should be taken into consideration and examined in Japanese, as well.

Third, from the point of view of animacy of noun phrases, the SRC in (5a) and the ORC in (5b) are different. It is well-known that the prototypical subject is animate whereas the prototypical object is inanimate (Comrie 1989). As a consequence, as we will show in the next section, there are much more instances of SRCs with an animate head-noun compared to those with an inanimate head-noun in corpora

(Sato 2011). On the contrary, there are more instances of ORCs with an inanimate head-noun compared to those with an animate head-noun. Since the head-noun is an animate noun phrase in (5), they are more expected in the SRC but less expected in the ORC. This might cause the difference in the processing.

Fourth, most of previous studies did not provide any discourse context in their experiments. The result of such experiments can be compared only if discourse contexts influence processing of SRC and ORC in the same way. There are studies, however, that point out the importance of discourse context in the processing of relative clauses (Mak, Vonk and Schfriers 2008; Roland et al. 2012). As it will be explained in detail in section 3, relative clauses play the important discourse function of introducing the referent into the discourse. For instance the ORC in (5b) sounds more natural if the preceding context introduces two reporters, one of them was criticized by the senator but the other was not. This indicates that not only syntactic or semantic factors, but also pragmatic factors should be taken into consideration.[6]

Finally, as we have pointed out in the introduction, differences between SRCs and ORCs are observed in the relative clause region in English but in the head-noun region in Japanese. It is true that, in either English or Japanese, relative clause processing involves some kind of association between the filler and the gap. They are, however, quite different in the direction of association; the filler comes first in English but the gap comes first in Japanese. See Kahraman et al. (2010), Kahraman (2011), Kwon et al. (2013) and Lin (2006) for detailed discussion. This might also be responsible for similarities and differences in the processing of English and Japanese relative clauses.

In summary, most of the above mentioned relative clause processing studies in Japanese did not pay enough attention to these various factors that might potentially influence relative clause processing. This is because most studies were conducted to test the validity of accounts based on theories on filler-gap dependency formation. In contrast, more and more researchers have recently begun to pay more attention to factors such as (un)predictability, frequency, animacy, discourse etc., which are closely related to statistical learning through experience. In the next section we will introduce studies examining such factors.

## 3 Relative clause processing and statistical learning

This section briefly introduces theories of human sentence processing based on theories that emphasize the role of statistical learning. They share the view that language processing mechanisms are shaped by learning through experience and that

**6** Even young English speaking children are sensitive to the presupposed contexts for relative clauses as shown in Hamburger and Crain (1982).

native speakers form probabilistic expectation for up-coming elements while they process sentences (Gennari and MacDonald 2009; MacDonald and Christiansen 2002; Wells et al. 2009). We first introduce theories that regard frequency or (un)predictability of certain structures or linguistic elements as responsible for determining processing difficulty. We then turn to theories that incorporate other types of factors, semantic factors such as animacy, or pragmatic factors such as discourse function. Then we will turn to recent studies in Japanese, which have attempted to examine the influence of various kinds of factors, such as case-marker driven expectation, frequency of occurrence, animacy, and discourse functions, in relative clause processing in Japanese.

## 3.1 Frequency of occurrence and (un)predictability for up-coming elements

Reali and Christiansen (2007)'s Frequency of Occurrence Hypothesis basically assumes that frequency of exposure to certain structures influences the difficulty of sentence processing. According to this hypothesis, highly frequent structures are processed more easily because readers/listeners are more familiar with frequent structures. Reali and Christiansen conducted a corpus study, and compared the frequencies of relative clauses in English. They found that SRCs have a higher frequency than ORCs. Then, they compared the noun types within relative clauses, and found that proper nouns are used more frequently in SRCs than ORCs, whereas pronouns are used more frequently in ORCs than SRCs. In other words, when the nouns within relative clauses are proper nouns, the frequency of a *that-verb-noun* chunk, like *that saw the teacher* was higher than a *that-noun-verb* chunk, like *that the teacher saw*. On the other hand, when pronouns were used within relative clauses, the frequency of a *that-pronoun-verb* chunk, like *that I/you/s(he) saw* was higher than a *that-verb-pronoun* chunk, like *that saw me/you/him/her*. Based on these findings, Reali and Christiansen manipulated the noun types within subject and object relative clauses, and conducted a series of self-paced reading experiments. The results showed that ORCs were read faster than SRCs when pronouns were used within relative clauses. Reali and Christiansen argued that co-occurrences of words, namely, the word-chunks play a central role in the sentence comprehension.

Recently, various theories have been proposed about the role of (un)predictability in human sentence processing, such as the Surprisal Hypothesis (Hale 2001; Levy 2008), and the *Entropy Reduction Hypothesis* (Hale 2006). Although these theories use slightly different metrics to estimate the (un)predictability of upcoming words or structures, the main logic is that unpredictability for upcoming elements determines the difficulty of sentence processing. According to the Surprisal Hypothesis (Hale 2001; Levy 2008), if the predictability of certain linguistic elements is lower than that of other elements, those elements are more difficult to process compared

to elements with higher predictability (Hale 2001; Levy 2008). Alternatively, Hale (2006) argues that reduction of uncertainty for up-coming elements increases the ease of sentence processing. Hale used a probabilistic context-free grammar and calculated the tree-bank probabilities of derivations for subject and object relative clauses, and found that probabilities of SRCs were higher than that of ORCs at relative pronoun *who*. This indicates that the uncertainty for up-coming elements was reduced in SRCs compared to ORCs at the relative pronoun position in English. In other words, English speakers expect more SRCs than ORCs at the relative pronoun positions. These theories, thus, attribute the processing difficulty of ORCs to their less predictable nature compared to SRCs.

Up to this point, we introduced accounts for relative clause processing based on the theories of frequency of occurrence or (un)predictability for up-coming elements. These accounts mainly focused on probabilistic distribution of linguistic elements classified with their lexical or syntactic properties. However, other researchers found that differences between SRCs and ORCs are not restricted to their lexical or syntactic properties. They proposed that semantic properties such as animacy of noun phrases or pragmatic properties such as discourse functions play important roles in relative clause processing. We will turn to those accounts in the next sub-section.

## 3.2 Semantic indeterminacy and expectation from discourse function

According to *Semantic Indeterminacy Hypothesis*, distributional patterns of animacy of noun phrases play a crucial role in forming probabilistic expectations in relative clause processing (Gennari and MacDonald 2008). Previous studies, which showed that SRCs were easier to process than ORCs, generally used animate nouns for the head-noun of relative clauses. However, various studies showed that when the head-noun of ORCs was an inanimate noun, ORCs were easier to process than SRCs (e.g., Mak, Vonk and Schfriers 2002, 2006; Traxler et al. 2005; Gennari and MacDonald 2008). Gennari and MacDonald (2008) pointed out that when the head-noun of relative clauses is an animate noun in an SRC, like *the reporter that attacked the senator admitted the error*, the thematic role of the head-noun (*the reporter*) is agent both for the relative clause verb (*attacked*) and the matrix verb (*admitted*). On the other hand, in the case of an ORC, like *the reporter that the senator attacked admitted the error*, the head-noun is patient of the relative clause verb and agent of the matrix verb. Gennari and MacDonald argued that the processing difficulty of ORCs may stem from the activation of several possible competing structures derived from the distributional pattern of thematic roles of given head-nouns. Moreover, they argued that ORCs with different animacy configurations may involve different competition processes between structural and semantic analyses, and this may cause a different amount of indeterminacy and processing difficulty. For example, Gennari

and MacDonald conducted a series of sentence completion experiments and found that when the head-noun of ORCs was an inanimate noun, it was most likely to be interpreted as the theme of the relative clause verb and agent of the matrix verb. On the other hand, when the head-noun of ORCs was an animate noun, it is likely to be interpreted as the agent, experiencer, patient or goal of the relative clause verb and theme of the matrix verb. These results suggest that when the head-noun of ORCs is an animate noun in English, there is a greater semantic indeterminacy for that noun, compared to inanimate noun.

Furthermore, Gennari and MacDonald, based on these findings, manipulated the head-noun animacy (animate vs. inanimate) and voice (active vs. passive) of ORCs, and conducted a self-paced reading experiment. The results showed that ORCs with inanimate head-nouns were read faster than ORCs with animate nouns. In other words, when the head-noun of ORCs' was inanimate, their processing was not as hard as that of ORCs with animate head-nouns. Based on these results, Gennari and MacDonald argued that processing difficulty of ORCs found in previous studies stems from the competition processes between structural and semantic analysis, which is closely related to distributional pattern of the head-noun animacy of relative clauses, and proposed *Semantic Indeterminacy Hypothesis*.

Fox and Thompson (1990) conducted a corpus study in English, and found that ORCs are generally used when the head-nouns are integrated (or 'grounded') into the ongoing discourse using old discourse referents. In Dutch, Mak, Vonk and Schfriers (2008) showed that when the noun phrase within relative clauses refers to the discourse topic in the previous context, the processing difficulty of ORCs is reduced. Based on the findings of Fox and Thompson (1990) and Mak, Vonk and Schfriers (2008), Roland et al. (2012) argued SRCs and ORCs are used for different purposes in real language. In order to extend and re-examine the findings of Fox and Thomson (1990), Roland et al. used various corpora and conducted a large scale corpus analysis in English. The results showed that an ORC, like *the student that the teacher saw…*, is more likely to appear in contexts in which the noun phrase within the relative clause has already been introduced as an old referent, like *the teacher was walking down the campus*. On the other hand, Roland et al. found that an SRC, like *the student saw the teacher…*is not likely to appear after such contexts. From these findings, Roland et al. pointed out that a noun phrase within an ORC is tends to be the topic of the ongoing discourse, whereas a noun phrase within an SRCs is not explicitly mentioned in the previous context, as argued in Fox and Thompson (1990). Roland et al. argued that, in terms of givenness of the noun phrases within relative clauses, a noun phrase within an ORC is an old discourse referent and a noun phrase within an SRC is a new referent. They then concluded that ORCs are generally used for grounding modified nouns to the ongoing discourse by using discourse old referent, whereas SRCs are used for supplying additional information about the modified noun by using new discourse referent. These observations suggest that SRCs and ORCs are used in different contexts for different purposes,

and ORCs are more dependent on the context than SRCs because ORCs are used in more specific situations (Fox and Thompson 1990).

Furthermore, Roland et al. argued that ORCs might have been more difficult to comprehend than SRCs because most of the previous studies examined the processing difficulty of these structures without any context. In other words, the lack of appropriate context might have violated the discourse requirements for ORCs, and this kind of unnaturalness might have caused the processing difficulty of ORCs. Moreover they pointed out that if the discourse requirements of ORCs are satisfied by an appropriate context and the unnaturalness is eliminated; the processing difficulty of ORCs might be reduced. In order to examine this hypothesis and provide empirical evidence, Roland et al. manipulated the contexts before the relative clauses as shown below, and compared the reading times of SRCs and ORCs through a series of self-paced reading experiments.

(7) a. Neutral context:
*There is always something happening in Elmwood Village.*

b. Topic context:
*The sculptor collected paintings.*

(8) a. SRC:
*The artist that admired the sculptor exhibited portraits at the gallery on Elmwood Avenue.*

b. ORC:
*The artist that the sculptor admired exhibited portraits at the gallery on Elmwood Avenue.*

In the Neutral context condition (7a), there is no particular topic, and no noun phrase within the relative clauses is explicitly mentioned. On the other hand, in the Topic context condition (7b), the embedded noun phrase within relative clauses, namely *the sculpture*, is subject and topic of the discourse. In a self-paced reading experiment, Roland et al. presented both SRCs and ORCs after the two types of contexts. The results showed that SRCs were read faster than ORCs in the Neutral context condition, as already shown in many previous studies in English (e.g., Staub 2010; see referents therein). In the Topic context condition, on the other hand, the reading times of SRCs and ORCs did not differ significantly. In other words, the processing difficulty of ORCs was reduced after the context like (7b), as in Dutch (Mak, Vonk and Schfriers 2008). Roland et al. argued that the results suggest that when the noun phrase within relative clauses is not mentioned in the previous context, and it is a new discourse referent, SRCs are easier to process than ORCs. On the other hand, when the noun phrase within relative clause is the topic of the previous

context, and it is an old discourse referent, ORCs are not harder to process than SRCs. Based on these results, Roland et al. concluded that the processing difficulty of ORCs stems from lack of appropriate context, and proposed the *Discourse Function Hypothesis*.

Overall, experience-based accounts suggest that in addition to structural complexity and memory load, probabilistic factors such as frequency, animacy, predictability, discourse function etc. are also strongly related to ease or difficulty of relative clause processing. In the remainder of this section, we will summarize the previous studies which have attempted to examine these factors in Japanese.

## 3.3 Relative clause processing in Japanese from the frequency of occurrence perspective

As explained above, previous studies have shown that frequency is one of the most important factors that influence the sentence processing of English (Reali and Christiansen 2007). In order to examine the possible effects of structural frequency on relative clause processing in Japanese, Sato, Kahraman and Sakai (2010a) conducted a corpus study. They analyzed the KOTONOHA Written Corpus (National Institute for Japanese Language and Linguistics). This corpus contains 10 million words. However, since it is not parsed either morphologically or syntactically, Sato, Kahraman and Sakai (2010a) used a 3 million word subset of the corpus, and they manually analyzed the distributions of SRCs and ORCs.

In total, Sato, Kahraman and Sakai (2010a) found 3187 relative clauses with transitive verbs. Of these, 1546 samples were tagged as SRCs, and 1641 samples were tagged as ORCs. This result demonstrated that the number of ORCs was slightly larger than SRCs. This indicates that the reading time patterns of SRCs and ORCs do not match with their distributions in the corpus because previous studies have consistently found SRCs were easier to process than ORCs in Japanese (Ishizuka 2005; Kahraman 2012; Kahraman et al. 2011a; Mitsugi, MacWhinney and Shirai 2010; Miyamoto and Nakamura 2003; Sakamoto and Yasunaga 2009; Ueno and Garnsey 2008). This result suggests that the processing difficulty of ORCs cannot be explained by its structural frequency alone in Japanese.

Apart from the study comparing the frequency of occurrence of SRCs and ORCs, there are other studies that have compared the frequencies of occurrence of other types of relative clauses. These studies also showed that the mere number of occurrences is not sufficient to explain the reading times. For example, Kim (2009) analyzed the distribution of so-called nominative-genitive conversion in Japanese. In Japanese, the subject noun phrase of adnominal clauses, including relative clauses, can be marked by a genitive case-marker as well as a nominative case marker, as shown in (9).

(9) *Gakusei ga/no kaita sakubun*
student NOM/GEN wrote composition
'the composition that the student wrote'

Kim counted the number of clauses with a nominative subject and that of clauses with a genitive subject, and found that adnominal clauses with genitive subjects are very limited in modern Japanese compared to those with nominative subjects.

In order to examine the possible effect of frequency in Japanese sentence processing, Kahraman (2012) also compared the reading times of ORCs with nominative and genitive nouns as shown below.

(10) a. Nominative-ORC:
*Sengetu gakusei ga kaita sakubun wa totemo*
last month student NOM wrote essay TOP very
*omosirokatta*
interesting
'The essay that the student wrote last month was very interesting.'

b. Genitive-ORC:
*Sengetu gakusei no kaita sakubun wa totemo*
last month student GEN wrote essay TOP very
*omosirokatta*
interesting
'The essay that the student wrote last month was very interesting.'

There is no clear meaning difference between (10a) and (10b). The only difference is the case marker of the subject noun within ORC. Kahraman (2012) assumed that if frequency is the decisive factor of processing difficulty in Japanese, the sentences in the *Genitive-ORC* condition should be read more slowly than in the *Nominative-ORC* condition. However the results showed that there was no significant difference between the two conditions. This result suggests that the frequency of occurrence alone is not sufficient to explain the differences in reading times. In addition to the structural frequency, this result also cannot be explained by the word-chunk frequency (c.f. Reali and Christianasen 2007). If the word-chunk frequency were a crucial factor in explaining processing difficulty, [NP-NOM RC-VERB] sequence would have been read faster than [NP-GEN RC-VERB] sequence, due to its higher word-chunk frequency[7].

7 Another instance comes from the study of the processing of cleft constructions in Japanese. Kahraman et al. (2011a) compared the reading times of subject and object clefts in a self-paced reading experiment. The results showed that the reading times of object clefts were faster than the reading times of subject clefts at the embedded verb position, suggesting that, unlike relative clauses, object clefts are easier to comprehend than subject clefts in Japanese. Furthermore, Kahraman et al. (2011b) conducted a corpus study, and analyzed the distribution of clefts in Japanese. The results showed that the frequency of subject clefts was significantly higher than that of object clefts. This finding suggests that the processing difficulty of subject clefts cannot be accounted for by their structural frequency in Japanese, as in the case of relative clauses.

As shown, the mere frequency of occurrence is not powerful enough to explain the processing difficulty in Japanese. However, this does not mean that frequency cannot account for the facts related to relative clause processing in Japanese. Rather, these studies suggest that frequency effects should be examined with a more fine-grained analysis of corpus distributions in Japanese. In this context, Sato, Kahraman and Sakai (2012) attempted to explore the role of animacy in the processing of Japanese SRCs and ORCs. Before we examine the role of animacy and revisit the influence of frequency, let us examine the role of probabilistic expectation derived from case-markers in Japanese.

## 3.4 Relative clause processing in Japanese from the case-driven expectation perspective

Previous studies have shown that Japanese speakers incrementally use the case-marker information, and start to build argument structure even before the verb is encountered (Kamide, Altmann and Haywood 2003; Miyamoto 2002; Yamashita 1997). According to these studies, the most important information source for the incremental sentence processing is case-markers in Japanese. For example, Kamide, Altmann and Haywood (2003) used a visual world paradigm and showed that as soon as Japanese speakers hear a dative noun soon after a nominative noun, they simultaneously start to look at the object, which can be an accusative noun. This study suggest that when Japanese speakers hear [NP-NOM > NP-DAT] sequence, they posit an accusative noun, and build [NP-NOM > NP-DAT > NP-ACC] argument structure even before the accusative noun is encountered. This indicates that case markers play an important role in Japanese sentence processing.

Based on the findings of these previous studies, Sato et al. (2009) pointed out that the case markers of the embedded nouns within SRCs and ORCs are different, and this difference may elicit different predictions for the upcoming elements and argument structure of the sentence (Ishizuka 2005). In Japanese, an SRC starts with an accusative noun, and an ORC starts with a nominative noun. Sato et al. (2009) argued that when Japanese native speakers see an accusative noun phrase initially, they immediately posit another noun, possibly an empty pronoun *pro*, because a sentence-initial nominative is missing (Miyamoto and Nakamura 2005). On the other hand, when Japanese native speakers see a sentence-initial nominative noun phrase, they do not predict a particular noun, as in the case of accusative noun because no element is missing in its canonical position. According to Sato et al. (2009), Japanese native speakers posit a missing noun when they encounter a transitive verb in an ORC. Therefore, the processing difficulty of ORCs may be caused by lack of prediction for another noun. The processing difficulty of ORCs over SRCs in Japanese can thus be accounted for by the *Case-driven Expectation Hypothesis*. According to this hypothesis, if there is an early expectation for another noun in

one structure, that structure should be easier to process than another structure where the expectation for another noun takes place later. In order to test this prediction, Sato et al. (2009) used relative clauses with various causative verb types, and conducted a series of sentence fragment completion experiments and self-paced reading experiments, as explained below.

In the sentence fragment completion task, Sato et al. first confirmed that when Japanese native speakers see an incomplete [NP-NOM > NP-DAT] sequence, like *sensei-ga gakusei-ni*, they add an accusative noun and then a verb, which indicates that an accusative noun is likely to be expected following the dative noun. On the other hand, when Japanese native speakers see an incomplete [NP-NOM > NP-ACC] sequence, like *sensei ga gakusei o*, they tend to add a verb, not a noun. This indicates that, there is no particular prediction for another noun when [NP-NOM > NP-ACC] sequence is encountered in Japanese. After confirming that [NP-NOM > NP-DAT] and [NP-NOM > NP-ACC] sequences elicit different expectations for the next word in Japanese, Sato et al. (2009) used relative clauses with causative verbs, in which the canonical word order of that causative structure is [NP-NOM > NP-DAT > NP-ACC] as shown in (11a) and (11b), and conducted a self-paced reading experiment.

(11) a. Dative-NP extracted-RC:

| | | | | | | |
|---|---|---|---|---|---|---|
| *Katyoo* | *ga* | *hisyo* | *o* | *kyooikusaseta* | *syain* | *wa* |
| manager | NOM | secretary | ACC | educated-CAUS | employee | TOP |

| | | |
|---|---|---|
| *syorui* | *o* | *nakusita.* |
| document | ACC | lost. |

‘The employee that the manager asked to educate the secretary lost the document.’

b. Accusative-NP extracted-RC:

| | | | | | | |
|---|---|---|---|---|---|---|
| *Katyoo* | *ga* | *hisyo* | *ni* | *kyooikusaseta* | *syain* | *wa* |
| manager | NOM | secretary | DAT | train-CAUS | employee | TOP |

| | | |
|---|---|---|
| *syorui* | *o* | *nakusita.* |
| document | ACC | lost. |

‘The employee that the manager asked the secretary to educate lost the document.’

In the Dative-NP extracted-RC condition, test sentences involve [NP-NOM > NP-ACC > RC-VERB > NP-TOP] sequence, and in the Accusative-NP extracted-RC condition, test sentences involve [NP-NOM > NP-DAT > RC-VERB > NP-TOP] sequence. Based on the results of sentence completion experiment, Sato et al. (2009) argued that an expectation for up-coming noun should be obtained earlier in the Accusative-NP extracted-RC condition than the Dative-NP extracted-RC condition because [NP-NOM > NP-DAT] sequence elicits an stronger expectation for an accusative noun, whereas [NP-NOM > NP-ACC] sequence does not elicit such an expectation. On the other

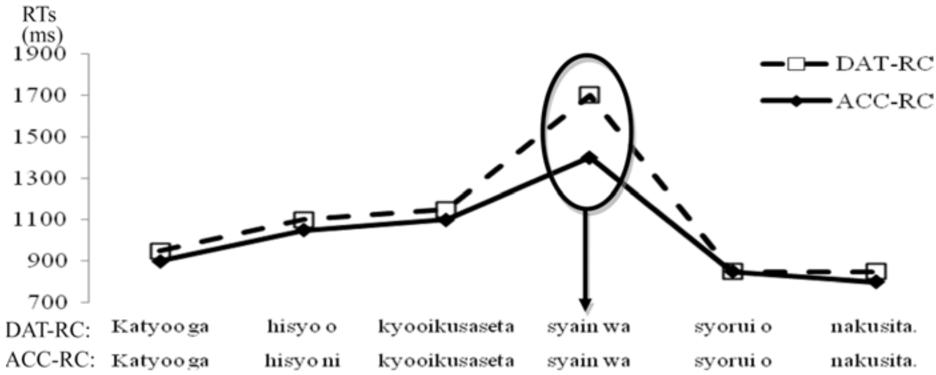



**Figure 4:** Reading times of Experiment 1 in Sato et al. (2009) (The circle and the arrow show where the significant processing asymmetry was observed.)

hand, structural distance between the head-noun and the gap is closer in the Dative-NP extracted-RCs than Accusative-NP extracted-RCs. Therefore, Sato et al. (2009) assumed that if the processing difficulty of ORCs in Japanese arises from the late expectation for another noun, the head-noun (*syain wa*) of the Accusative-NP extracted-RC should be read faster than that of Dative-NP extracted-RC due to an early expectation for another noun. On the other hand, Sato et al. (2009) pointed out that if the structural distance between the head-noun and the gap is a more important factor in the processing of Japanese relative clauses, the head-noun of Dative-NP extracted-RC should be read faster than that of Accusative-NP extracted-RC.

The results of a self-paced reading experiment showed that the head-noun of relative clauses was read faster in the Accusative-NP extracted-RC condition than the Dative-NP extracted-RC condition. This result is not consistent with the predictions of *Structural Distance Hypotheses* (Hawkins 1999, 2004; O'Grady 1997). On the other hand, this result is in line with the theory proposed by Sato et al. (2009), which predicts that an early expectation for another noun facilitates the processing of the relative clause in Japanese.

Although this result supports the hypothesis assumed by Sato et al., it is still consistent with the predictions of the *Dependency Locality Theory* (Gibson 1998, 2000). Since it was hard to distinguish between the effect of early expectation for another noun and linear distance between the head-noun and the gaps, Sato et al. (2009) conducted another experiment. In this experiment, Sato et al. again used relative clauses with causative verbs. However, this time, unlike in the first experiment, they used causative sentences with a verb such as *dookoo-suru* 'join' that takes a dative argument. The causative sentence derived from such verbs, namely, (12) has the canonical word order [NP-NOM > NP-ACC > NP-DAT], in which the accusative noun precedes the dative noun. See Kuno (1973) and Shibatani (1972) for detailed explanations for the word orders of causative structures in Japanese.

(12)

| | | | | | |
|---|---|---|---|---|---|
| *Butyoo* | *ga* | *kakarityoo* | *o* | *sinnyuusyain* | *ni* |
| general manager | NOM | subsection chief | ACC | new employee | DAT |

*dookoosaseta.*
accompany-CAUS
'The general manager made the subsection chief join the new employee.'

The linear distance between the head-noun and gap position is greater in the Accusative-NP extracted-RC condition. Therefore, the Dependency Locality Theory predicts that the head-noun of the relative clause should be read faster in the Dative-NP extracted-RC condition than the Accusative-NP extracted-RC condition. On the other hand, Sato et al. (2009)'s case-driven expectation theory predicts a reverse pattern.[8]

(13) a. Dative-NP extracted-RC:

| | | | | | |
|---|---|---|---|---|---|
| *Butyoo* | *ga* | *kiremono* | *no* | *kakarityoo* | *o* |
| general manager | NOM | clever | GEN | subsection chief | ACC |
| *dookoosaseta* | *sinnyuusyain* | *wa* | *syorui* | *o* | *wasureta.* |
| accompany-CAUS | new employee | TOP | document | ACC | forgot |

'The new employee that the general manager asked the smart subsection chief to accompany forgot the document.'

b. Accusative-NP extracted-RC:

| | | | | | |
|---|---|---|---|---|---|
| *Butyoo* | *ga* | *kiremono* | *no* | *kakarityoo* | *ni* |
| general manager | NOM | clever | GEN | subsection chief | DAT |
| *dookoosaseta* | *sinnyuusyain* | *wa* | *syorui* | *o* | *wasureta.* |
| accompany-CAUS | new employee | TOP | document | ACC | forgot |

'The new employee that the general manager asked to accompany the smart subsection chief forgot the document.'

The results of the self-paced reading experiment again showed that the head-nouns of relative clauses were read faster in the Accusative-NP extracted-RC condition than the Dative-NP extracted-RC condition. This result is not consistent with the predictions of the *Dependency Locality Theory* (Gibson 1998, 2000). On the other hand, this result is consistent with the prediction by Sato et al. (2009).

**8** The original test sentences in Sato et al. (2009) were simplified due to space limitations in Figure 5. The original sentences were *sihatu no densya de butyoo ga kiremono no kakarityoo o dookoosaseta sinnyuusyain wa syorui o wasureta* ('the new employee that the general manager asked the smart subsection chief to accompany on the first train forgot the document') and *sihatu no densya de butyoo ga kiremono no kakarityoo ni dookoosaseta sinnyuusyain wa syorui o wasureta* ('the new employee that the general manager asked to accompany the smart subsection chief on the first train forgot the document'). Sato et al. (2009) reported no significant difference in reading time at the region of *sihatu no densya de* ('on the first train').

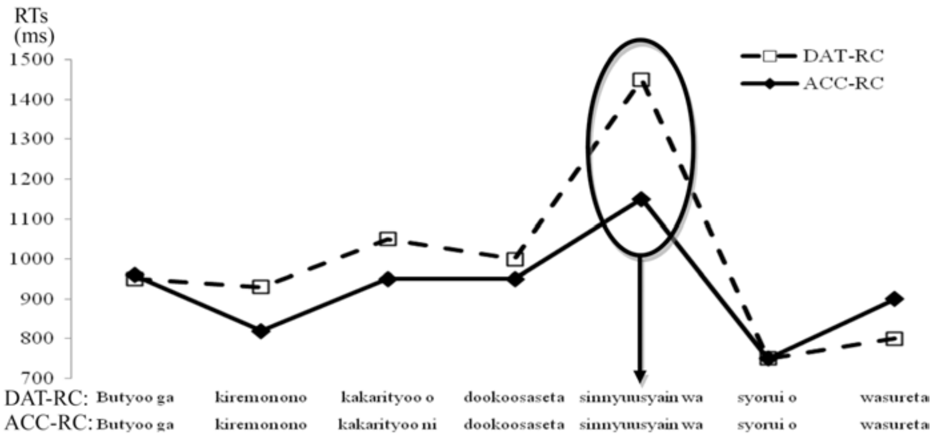



**Figure 5:** Reading times of Experiment 2 in Sato et al. (2009) (The circle and the arrow show where the significant processing asymmetry was observed.)

Taken together, Sato et al. (2009) showed that neither the *Structural Distance Hypothesis* nor the *Dependency Locality Theory* can explain the entire set of results. On the other hand, the results of these two experiments are consistent with the predictions derived from their hypothesis. They assumed that if there is an early expectation for a noun in a particular structure, that structure is processed more easily compared with the other structure, in which the expectation for another noun takes place later. According to this account, case markers play a central role in making the prediction of argument structure and upcoming elements in Japanese. If the case markers immediately signal the existence of another noun, the later part of that sentence is processed easily due to early and higher expectation for the upcoming noun. On the other hand, if a case marker, like sentence initial nominative case marker in ORCs, does not directly signal the existence of another noun, that structure is more difficult to process, due to late and less activated expectation for the up-coming noun.

The essential metric for the *Case-driven Expectation* account is quite different from predictability based accounts such as the *Surprisal Hypothesis* (Hale 2001; Levy 2008) or the *Entropy Reduction Hypothesis* (Hale 2006). Remember that we called the latter two hypotheses (Un)predictability Hypotheses since they both assumed that occurrence of unpredictable structures or uncertainty of up-coming elements increase processing difficulty. Thus, the Case-driven Expectation Hypothesis cannot be directly adapted to other languages that lack an explicit case marking system. However, Sato et al. (2009) showed that basic ideas of accounts based on the (Un)predictability Hypotheses can also be effective in the relative clause processing research on Japanese. In addition, this study importantly suggests that the structure-based accounts such as the Dependency Locality Theory or the Structural Distance Hypothesis cannot account for the whole range of facts observed in Japanese relative

clause processing studies, and other factors, such as case-driven expectation should be taken into consideration.

Finally, as we have already shown, the Case-driven Expectation Hypothesis can account for the asymmetry between SRCs and ORCs as well as the asymmetry between the Accusative-NP extracted-RC and the Dative-NP extracted-RC. Taken together, the Case-driven Expectation Hypothesis accounts for the fact that the ease of relative clause processing is in accordance with the hierarchy of nominative > accusative > dative, as proposed by the *Noun Phrase Accessibility Hierarchy* by Keenan and Comrie (1977). The Case-driven Expectation Hypothesis is thus enables us to integrate the insight of typological studies into the psycholinguistic studies of relative clause processing.

## 3.5 Relative clause processing in Japanese from the semantic indeterminacy perspective

The ORCs' structural frequency is slightly higher than SRCs in Japanese. However, this distributional pattern does not match with their reading times. Previous studies in English and Dutch showed that when the head-noun of ORCs was an inanimate noun, they were easier to process than process than SRCs (e.g., Mak, Vonk and Schfriers 2002, 2006; Traxler et al. 2005). Gennari and MacDonald (2008) pointed out that these results are likely to be due to semantic indeterminacy, which is closely related to the distributional patterns of animacy in relative clauses. According to Gennari and MacDonald, ORCs are more likely to occur with inanimate head-nouns, whereas SRCs occur with animate head-nouns.

Based on these studies, Sato, Kahraman and Sakai (2012) analyzed the distribution of animate and inanimate nouns in relative clauses through a corpus study. The results showed that the animacy of head-nouns of SRCs and ORCs differ in Japanese. In SRCs, the head-noun is generally an animate noun. On the other hand, the head-noun of the ORCs is an inanimate noun in most cases. In short, they observed the same tendency in Japanese as Gennari and MacDonald (2008) observed in English. In addition to Japanese corpus analysis, Sato, Kahraman and Sakai also conducted a sentence completion experiment in Japanese. The results were in line with the distributional patterns of animacy in relative clauses. Participants produced SRCs with animate head-nouns, whereas the head-nouns of ORCs were inanimate nouns generally.

Based on these results, Sato, Kahraman and Sakai argued that the processing difficulty of ORCs can be explained by frequency once we take the animacy of the head-noun into consideration. Most of the previous studies used animate head-nouns in the experiments. Since the distributional pattern of the animacy of head-noun favored SRCs, they might have been found easier to process in Japanese. In order to examine this possibility they manipulated the animacy of SRCs and ORCs, as shown below, and conducted the self-paced reading experiment.

(14) a. SRC:

| | | | | | | |
|---|---|---|---|---|---|---|
| *Kokkai* | *no* | *honkaigi* | *de* | *syoogen* | *o* | *kobanda* |
| Diet | GEN | meeting | LOC | testimony | ACC | refused |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *giin* | *wa* | *izen* | *kara* | *tyuumoku* | *no* | *mato* | *datta.* |
| senator | TOP | before | *from* | attention | GEN | center | COP |

'The senator that refused to testify at the National Diet meeting was the center of attention for a long time.'

b. ORC:

| | | | | | | |
|---|---|---|---|---|---|---|
| *Kokkai* | *no* | *honkaigi* | *de* | *giin* | *ga* | *kobanda* |
| Diet | GEN | meeting | LOC | senator | NOM | refused |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *syoogen* | *wa* | *izen* | *kara* | *tyuumoku* | *no* | *mato* | *datta* |
| testimony | TOP | before | from | attention | GEN | center | COP |

'The testimony that the senator refused at the National Diet meeting was the center of attention for a long time.'

In the SRC condition, the embedded noun phrase within the relative clause is an inanimate noun, and the head-noun is an animate noun. Conversely, in the ORC condition, the embedded noun phrase within relative clause is an animate noun, and the head-noun is an inanimate noun. The length and frequency of head-nouns are carefully matched between the conditions. Since there are more ORCs compared to SRCs in the corpus as reported by Sato et al. (2010a), Sato, Kahraman and Sakai (2012) argued that if the processing difficulty of ORCs stems from the head-noun animacy, ORCs may be processed more easily than SRCs. The results are shown in Figure 6.

The results of a self-paced reading indeed showed that the head-noun of the ORC condition (*syoogen-wa*) was read faster than *giin-wa* in the SRC condition. This result suggests that when the inanimate nouns were used as the head-nouns of

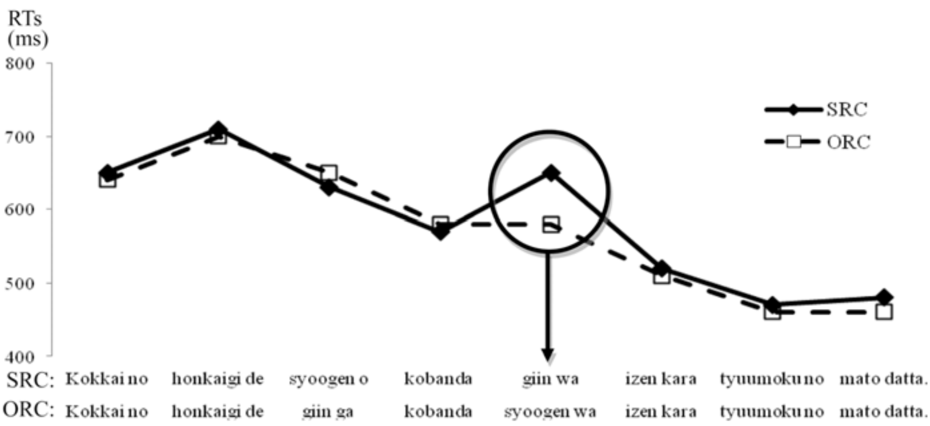



**Figure 6:** Reading times in Sato, Kahraman and Sakai (2012) (The circle and the arrow show where the significant processing asymmetry was observed.)

ORCs, they were easier to process than SRCs in Japanese. In addition to this study, Bai and Hirose (2013) and Nakamura and Miyamoto (2013) also reported similar results, and showed that ORCs with an inanimate head-noun were read faster than SRCs with an animate head-nouns in Japanese. The results also suggest that in addition to syntactic information from case markers, native speakers of Japanese also incrementally use the semantic information from the animacy of noun phrases. Moreover, these studies suggest that distributional patterns of noun phrases within relative clauses might also play an important role beyond that played by structural frequencies in processing of Japanese relative clauses.

## 3.6 Relative clause processing in Japanese from the discourse function perspective

We have seen that both syntactic information from case markers and semantic information from animacy are used in the processing of relative clauses in Japanese. Here, we will introduce one more study which deals with the influence of pragmatic factors, namely discourse function, on processing of Japanese relative clauses.

As we explained above, Roland et al. (2012) showed that discourse functions of SRCs and ORCs differ, and when an appropriate context is provided, ORCs are no longer harder to process than SRCs. In other words, according to Roland et al. (2012), processing difficulty of ORCs arises from the lack of appropriate context, which violates the discourse requirements for ORCs.

In order to examine the influence of discourse function in Japanese, Sato, Kahraman and Sakai (2010b) conducted a corpus study and self-paced reading experiments.[9] In the corpus study, they investigated the newness of the discourse referents within relative clauses. They analyzed the sentences before SRCs and ORCs within 400 characters including *kanji* and *kana*. The results showed that 70% of the noun phrases within SRCs are new discourse referents, whereas 80% of the noun phrases within ORCs are old discourse referents. This result shows that the embedded noun phrases within SRCs are not necessarily mentioned in the preceding context on the one hand. On the other hand, the embedded noun phrases within ORCs are explicitly mentioned in the preceding context in Japanese as observed in English (Roland et al. 2012). This finding suggests that the contexts prior to RCs are very similar in English and Japanese. Based on this result, Sato, Kahraman and Sakai (2010b) hypothesized that if the discourse functions of relative clauses are universal, the processing difficulty of ORCs would be reduced, when an appropriate context provided in Japanese. In order to examine this hypothesis, they prepared topic and neutral contexts before relative clauses, as shown below, and conducted a self-paced reading experiment.

9 Sato et al. (2010b) cites the earlier versions of Roland et al. (2012) study.

(15) Neutral Context

| | | | | | | |
|---|---|---|---|---|---|---|
| *Insanna* | *satuzinziken* | *no* | *genba* | *de* | *soosa* | *ga* |
| gruesome | murder case | GEN | place | LOC | investigation | NOM |

| |
|---|
| *okonawareta.* |
| was.carried.out. |

'An investigation was carried out at the site of the gruesome murder case.'

a. SRC

| | | | | |
|---|---|---|---|---|
| *Tokusoobu* | *no* | *keizi* | *o* | *yobitometa* |
| special investigation department | GEN | detective | ACC | called |

| | | | | | |
|---|---|---|---|---|---|
| *tantoosya* | wa | *temizikani* | *genba* | *o* | *annaisita.* |
| officer | TOP | quickly | place | ACC | introduced |

'The officer who called the special investigation department detective quickly introduced the site of the murder case.'

b. ORC

| | | | | |
|---|---|---|---|---|
| *Tokusoobu* | *no* | *keizi* | *ga* | *yobitometa* |
| special investigation department | GEN | detective | NOM | called |

| | | | | | |
|---|---|---|---|---|---|
| *tantoosya* | *wa* | *temizikani* | *genba* | *o* | *annaisita.* |
| officer | TOP | quickly | place | ACC | introduced |

'The officer who the special investigation department detective called quickly introduced the place of murder case.'

(16) Topic Context

| | | | | | |
|---|---|---|---|---|---|
| *Tokusoobu* | *no* | *keizi* | *ga* | *ziken* | *no* |
| special investigation department | GEN | detective | NOM | incident | GEN |

| | | |
|---|---|---|
| *soosa* | *ni* | *atatta.* |
| investigation | DAT | assigned |

'The criminal prosecutor was assigned to the investigation of an incident.'

a. SRC

| | | | | | | |
|---|---|---|---|---|---|---|
| *Sono* | *keizi* | *o* | *yobitometa* | *tantoosya* | *wa* | *temizikani* |
| that | detective | ACC | called | officer | TOP | quickly |

| | | |
|---|---|---|
| *genba* | *o* | *annaisita.* |
| place | ACC | introduced |

'The officer who called that detective quickly introduced the place of incident.'

b. ORC

| | | | | | | |
|---|---|---|---|---|---|---|
| *Sono* | *keizi* | *ga* | *yobitometa* | *tantoosya* | *wa* | *temizikani* |
| that | detective | NOM | called | officer | TOP | quickly |

| | | |
|---|---|---|
| *genba* | *o* | *annaisita.* |
| place | ACC | introduced |

'The officer who that detective called quickly introduced the place of incident.'

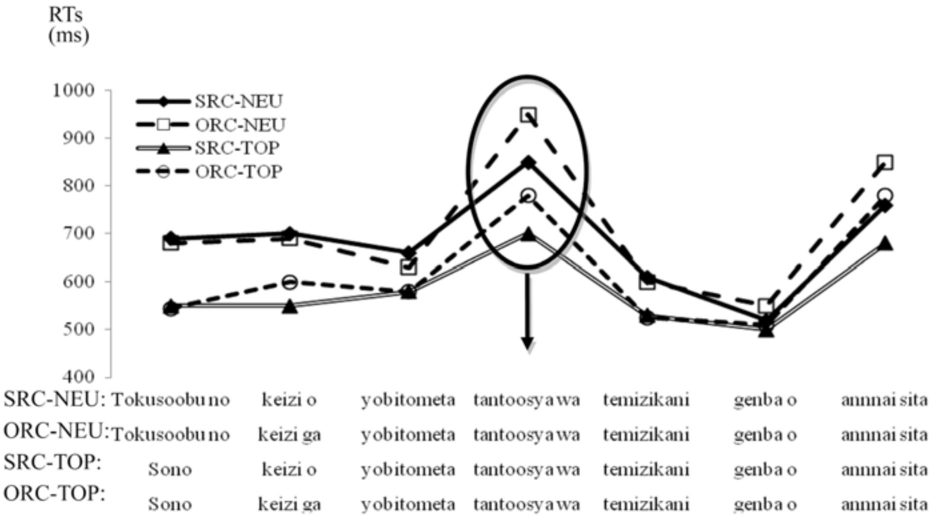



**Figure 7:** Reading times of Sato, Kahraman and Sakai (2010b) (The circle and the arrow show where the significant processing asymmetry was observed.)

In the Neutral Context condition (15), the embedded noun phrase within relative clauses (*keizi*) is not overtly mentioned in the preceding context. On the other hand, in the Topic Context condition (16), the embedded noun phrase within relative clauses (*keizi*) is overtly mentioned, and topic of the preceding context.

The results of a self-paced reading experiment showed that the head-noun (*tantoosya-wa*) of SRCs was read faster than the head-noun of ORCs in the both Topic and Neutral context conditions as shown in Figure 7. At first glance, this result seems to show that, unlike English, SRCs were easier to process than ORCs in both contexts in Japanese, and suggests that the processing difficulty of ORCs in Japanese cannot be accounted for by the Discourse Function Hypothesis.

However, there is a crucial difference between the English relative clause and the Japanese relative clause. Since the relative clause follows the head-noun in English, the previous studies in English compared the processing time in the relative clause region. In Japanese, since the relative clause precedes the head-noun, most of studies compare the reading time in the head-noun region. Importantly, the previously mentioned noun phrase appears in the critical region in English. In Japanese, on the other hand, the previously mentioned noun phrase does not appear in the critical region. The occurrence of the previously mentioned noun is crucial in the discourse function account because the head-nouns are integrated into the context using those ‘old discourse referents’. Therefore, Sato, Kahraman and Sakai (2010b) suggest that the previous context influences the processing difficulty of relative clauses only if we compare the regions that contain the previously mentioned noun phrase. The result of their experiment thus suggests that, even though the discourse

functions of relative clauses are universal, the typological differences between English and Japanese might obscure the processing asymmetry caused by discourse functions.

## 4 Conclusion and future direction

This chapter focused on the online processing of relative clauses in Japanese, and reviewed previous studies that examined various factors that might influence relative clause processing. Earlier studies tended to limit their attention to structural factors such as constituent order or structural complexity, whereas more recent studies gradually shifted their attention to more wide range of factors, such as expectation, frequency, animacy of noun phrases, and discourse function of relative clauses. These studies tell us that the processing difficulty of relative clauses is not determined just by a single factor, and many different factors influence the processing of Japanese relative clauses. These studies, at the same time, revealed that various kinds of typological differences among languages impact sentence processing in real time.

Section 1 introduced a processing asymmetry between subject relative clauses (SRCs) and object relative clauses (ORCs). Generally, SRCs are easier to process compared to ORCs in most languages, irrespective of their typological differences. We pointed out that various accounts have been proposed to explain the processing asymmetry between SRCs and ORCs. They are roughly classified into two groups; the accounts based on the theories related to filler-gap dependency formation and the accounts based on the processing theories emphasizing the importance of statistical learning through experience. Then, we explained that most of the existing accounts have been proposed based on data from English. Since they make similar predictions for the processing difficulty of relative clauses, examination of the source of processing asymmetry between SRCs and ORCs in typologically different languages such as English and Japanese would provide important insights into characterization of human sentence processing mechanism.

In section 2, we first briefly explained the accounts based on the theories on filler-gap dependency formation: Dependency Locality Theory (Gibson 1998, 2000) and Structural Distance Hypothesis (Hawkins 1999, 2004; O'Grady 1997). We showed that both accounts predict that SRCs should be easier to process than ORCs in English, whereas the predictions of these accounts differ in Japanese. Relative clause processing studies in Japanese was to examine the validity and universality of these accounts, and they found SRCs were easier to process than ORCs in Japanese, and these results support the Structural Distance Hypothesis over the Dependency Locality Theory.

Section 3 briefed the representative instances of accounts emphasizing the role of learning through experience, such as *Frequency of Occurrence Hypothesis* (Reali and Christiansen 2007), *Predictability (or Reduction of Uncertanity) Hypotheses* (Hale 2001, 2006, Levy 2008), *Semantic Indeterminacy Hypothesis* (Gennari and MacDonald, 2008) and *Discourse Function Hypothesis* (Roland et al. 2012). We then reviewed previous studies that have examined Japanese relative clause processing from the viewpoint of frequency of occurrence. We first reviewed studies by Sato, Kahraman and Sakai (2010a) and Kahraman (2012). They showed that the distributions and reading times' patterns do not match in Japanese. It is therefore difficult to explain the processing difficulty just by the structure or word-chunk frequency in Japanese.

Sato et al. (2009), using causative relative clauses, examined the impact of case-driven expectation in Japanese. As a language-specific extension of the (Un) predictability Hypotheses in Japanese, their study suggested that expectations for upcoming or missing nouns derived from case markers play very crucial role in the processing of relative clauses in Japanese.

Sato, Kahraman and Sakai (2012) attempted to examine the animacy effects in Japanese relative clause processing. This study showed that SRCs tend to occur with animate head-noun, whereas ORCs tend to occur with inanimate head-nouns in Japanese. Moreover, this study also showed that ORCs with inanimate head-nouns were easier to process than SRCs with animate head-nouns in Japanese. These results also suggested that in addition to syntactic information obtained from case markers, semantic information from the animacy of noun phrases is also used very effectively in relative clause processing, and distributional patterns of noun phrases within relative clauses is an important factor in Japanese.

Sato, Kahraman and Sakai (2010b) dealt with the examination of the impact of discourse function on relative clause processing in Japanese. The results of this study showed noun phrases within SRCs are not explicitly mentioned in the preceding context, whereas the noun phrases within ORCs tend to be mentioned explicitly in the previous context in Japanese, as in English (Roland et al. 2012). Moreover, the results of self-paced reading experiments showed that the type of previous context has also an impact on the processing of relative clauses. However, the head-noun of ORCs was still harder to process than SRCs in both neutral and topic contexts, indicating that the processing difficulty of ORCs itself does not from the lack of appropriate context in Japanese, unlike English.

Overall, these studies clearly suggest that the processing ease or difficulty of relative clauses cannot be reduced to a single factor in Japanese. They have narrowed down the possible factors that may affect relative clause processing in Japanese. Nevertheless, it is by no means true that these factors are sufficient to explain relative clause processing in Japanese. Many more studies and empirical evidence are needed to explain the processing ease or difficulty of relative clauses. In addition, as we pointed out above, various factors seem to simultaneously affect relative clause processing. Therefore, the examination of interaction between various

factors such as case-driven expectation and animacy, or discourse function and animacy etc. would be important.

In the reminder of this chapter, we will speculate about future directions of relative clause processing studies in Japanese. First of all, close collaboration between psycholinguistic research on human sentence processing and computational linguistic research on large scale corpora is necessary. We explained that various kinds of probabilistic factors play important role in relative clause processing. For instance, we showed that structural frequency alone cannot explain the difficulty of relative clause processing in Japanese (Kahraman 2012, Kahraman et al. 2011b; Sato, Kahraman and Sakai 2010a), whereas the distributional patterns of head-noun animacy can provide more accurate explanation for the processing difficulty (Sato, Kahraman and Sakai 2012). This result strongly suggests that researchers need to investigate various kinds of distributional patterns and should not be satisfied with just structural frequencies. Further corpus analyses of more factors are necessary to explain the experimental results.

Finally, according to Jaeger and Norcliffe (2009), cross-linguistic studies on typologically diverse languages in the world are still very limited. Research on Japanese relative clause processing can contribute to the psycholinguistic studies from such perspective. Notice that debates continue in the field of typological studies on relative clauses (see, Comrie 1998, 2010, among others). According to Comrie, relative clauses are classified into *European-Type* and *Asian-Type*. The European-Type relative clauses are overtly marked by relative clause markers and involve clear instances of filler-gap dependency whereas the Asian-Type relative clauses lack either of them. Psycholinguistic researchers have so far concentrated on European-Type relative clauses. However, the Japanese language, for instance, allows *gapless relative clauses* such as (17a) (see Yamashita, Stowe and Nakayama 1993), and *head-internal relative clauses* such as (17b).

(17) a.

| | | | |
|---|---|---|---|
| *Sakana* | *ga* | *yakeru* | *nioi* |
| fish | NOM | to bake | smell |

‘The smell of the fish that is baked.’

b.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Doroboo* | *ga* | *ie* | *ni* | *haitta* | *no* | *o* | *tukamaeta.* |
| thief | NOM | house | DAT | got into | COMP | ACC | caught |

‘(I /somebody) caught the thief who got into house.’

Other researchers argue that these constructions are actually sentential modifier/adnominal clauses and it might be problematic to group them together with typical relative clauses with a filler and a gap. Nevertheless, the investigation of the processing of such constructions would provide valuable insight into the universal cognitive mechanisms for human sentence processing and shed light onto the debate on the typological classification of relative clauses.

We hope that psycholinguistic research on Japanese relative clause processing will develop further in the future and provide us with the answers for the fundamental questions in the cognitive science of language: What is the nature of human cognitive mechanisms for language processing? How can our cognitive mechanisms process typologically diverse languages in the world?

## Acknowledgments

We would like to express our sincere thanks to Atsushi Sato, Hajime Ono, and Douglas Roland. Atsushi is in fact the 'hidden third author' of this chapter. Without his profound knowledge and extraordinary passion in conducting research on Japanese relative clause processing, the content of this chapter would have come out to be less interesting and exciting. Hajime helped us at the starting point of relative clause processing research in our lab and participated in many of the subsequent studies reported in this chapter. Doug kindly gave us valuable comments and suggestions on the earlier version of this chapter. We would cordially thank the editor of this volume, Mineharu Nakayama for his kindness. Without his extraordinary patience and generosity, this chapter could not have been included in this volume. Also we would like to thank the reviewers for their constructive comments. Parts of the research reported in this chapter were supported by a grant-in-aid for scientific research (A) from the Japanese Society for the Promotion of Science (Typological variation and plasticity of neurocognitive mechanism for language processing: A view from comparative studies on East Asian Languages; P.I., Hiromu Sakai, Hiroshima University). We also acknowledge the Japanese Society for the Promotion of Science for the support to the first author (Sentence processing of Japanese and Turkish as a first and second language; Project No: 25/03004) All remaining shortcomings are, of course, our own.

## References

Bai, Chunhua and Yuki Hirose. 2013. Nihongo ni okeru shugo-mokutekigo kankeisetsu shori no kosuto no saikentō [A revisitation of the processing cost of Japanese subject and object relative clauses]. *IECIE Technical Report* 113(174). 27–32.

Betancort, Moises, Manuel Carreiras and Patrick Sturt. 2009. The processing of subject and object relative clauses in Spanish: An eye-tracking study. *Quarterly Journal of Experimental Psychology* 62(10). 1915–1929.

Carreiras, Manuel, Jon Andoni Duñabeitia, Marta Vergara, Irene de la Cruz-Pavía and Itziar Laka (2010). Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition* 115(1) .79–92.

Cohen, Laurent and Jacques Mehler. 1996. Click monitoring revisited: An on-line study of sentence comprehension. *Memory and Cognition* 24(1). 94–102.

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology.* Chicago, IL: University of Chicago Press.

Comrie, Bernard. 1998. Rethinking the typology of relative clauses. *Language Design* 1. 59–86.
Comrie, Bernard. 2010. Japanese and the other languages of the world. NINJAL Project Review 1. 29–45.
Fodor, Janet D. 1989. Empty categories in sentence processing. *Language and Cognitive Processing* 4(3–4). 155–209.
Fox, Barbara A. and Sandra A. Thompson. 1990. A discourse explanation of the grammar of relative clauses in English conversation. *Language* 66(2). 297–316.
Gennari, Silvia P. and Maryellen C. MacDonald 2008. Semantic indeterminacy in object relative clauses. *Journal of Memory and Language* 58(2). 161–187.
Gennari, Silvia P. and Maryellen C. MacDonald 2009. Linking production and comprehension process: The case of relative clauses, *Cognition* 111(1), 1–23.
Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1), 1–76.
Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Alec P. Marantz, Yasushi Miyashita and Wayne O'Neil (eds.), *Image, language brain: Papers from the first Mind Articulation Project Symposium*, 95–126. Cambridge, MA: MIT Press.
Gordon, Peter C., Randall Hendrick and Marcus Johnson. 2001. Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27(6). 1411–1423.
Hale, John. 2001. A probabilistic early parser as a psycholinguistic model. *Proceedings of North American Association for Computational Linguistics*, Vol. 2. 159–166.
Hale, John. 2006. Uncertainty about the rest of the sentence. *Cognitive Science* 30(4). 609–642.
Hamburger, Henry and Stephen Crain. 1982. Relative acquisition. In Stan Kuczaj (ed.), *Language development Vol. 1: Syntax and semantics*, 245–274. Hillsdale, NJ: Erlbaum.
Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75(4). 244–285.
Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
Hsiao, Franny and Edward Gibson. 2003. Processing relative clauses in Chinese. *Cognition* 90(1). 3–27.
Ishizuka, Tomoko. 2005. Processing relative clauses in Japanese. *UCLA Working Papers in Linguistics* 13(2). 135–157.
Jaeger, Florian T. and Elisabeth J. Norcliffe. 2009. The cross–linguistic study of sentence production: State of the art and a call for action. *Language and Linguistic Compass* 3(4). 866–887.
Kahraman, Barış. 2011. *Nihongo oyobi Torukogo ni okeru "kūsho to maigo no izon kankei" no shori: Bunshori no chikujisei o megutte* [Processing "gap-filler dependencies" in Japanese and Turkish: Regarding the incrementality of sentence processing]. Hiroshima, Japan: Hiroshima University dissertation.
Kahraman, Barış. 2012. Sentence processing of nominative-genitive conversion in Japanese by Turkish speaking learners and native speakers. In Yukio Otsu (ed.), *The Proceedings of Thirteenth Tokyo Conference on Psycholinguistics*. 81–102. Tokyo: Hituzi Syobo.
Kahraman, Barış, Atsushi Sato, Hajime Ono and Hiromu Sakai. 2010. Relative clauses processing before the head-noun: Evidence for strong forward prediction in Turkish. In Hiroki Maezawa and Azusa Yokogoshi (eds.), *MIT Working Papers in Linguistics 61*, 155–170. Cambridge, MA: MITWPL.
Kahraman, Barış, Atsushi Sato, Hajime Ono and Hiromu Sakai. 2011a. Incremental processing of gap-filler dependencies: Evidence from the processing of subject and object clefts in Japanese. In Yukio Otsu (ed.), *The Proceedings of Twelfth Tokyo Conference on Psycholinguistics*. 133–147. Tokyo: Hituzi Syobo.
Kahraman, Barış, Atsushi Sato, Hajime Ono and Hiromu Sakai. 2011b. Why object clefts are easier to process than subject clefts in Japanese: Frequency or expectation?. *IECIE Technical Report* 111(170). 67–72.
Kamide, Yuki, Gerry T. M. Altmann and Sarah L. Haywood. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye-movements. *Journal of Memory and Language* 49(1). 133–156.

Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 7(1). 63–99.

Kim, Eunju. 2009 Gendaigo-no rentai-shūshokusetsu-ni okeru joshi “no”. [The particle ‘no’ in the modificational construction in modern Japanese]. *Nihongo Kagaku* 25. 23–42.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT press.

Kwon, Nayoung, Robert Kluender, Marta Kutas and Maria Polinsky. 2013. Subject/object processing asymmetries in Korean relative clauses: Evidence from ERP data. *Language* 89(3). 537–585.

Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition* 106(3). 1126–1177.

Lin, Chien-Jer Charles. 2006. *Grammar and parsing: A typological investigation of relative-clause processing*. Tucson, AZ: The University of Arizona dissertation.

Lin, Chien-Jer Charles. 2014. Effect of thematic order on the comprehension of Chinese relative clauses. *Lingua* 140. 180–206.

MacDonald, Maryellen C. and Morten H. Christiansen. 2002. Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review* 109 (1). 35–54.

MacWhinney, Brian and Csaba Pleh. 1988. The processing of restrictive relative clauses in Hungarian. *Cognition* 29(2). 95–141.

Mak, Willem M., Wietske Vonk and Herbert Schfriers. 2002. The influence of animacy on relative clause processing. *Journal of Memory and Language* 47(1). 50–68.

Mak, Willem M., Wietske Vonk and Herbert Schfriers. 2006. Animacy in processing relative clauses: The hikers that rocks crash. *Journal of Memory and Language* 54(4). 466–490.

Mak, Willem M., Wietske Vonk and Herbert Schfriers. 2008. Discourse structure and relative clause processing. *Memory & Cognition* 36(1). 170–181.

Mitsugi, Sanako, Brian MacWhinney and Yasuhiro Shirai. 2010. Cue-based processing of relative clauses in L2 Japanese. In Matthew T. Prior, Yukiko Watanabe, and Sang-Ki Lee (eds.), *Selected Proceedings of the 2008 Second Language Forum*, 123–138. Somerville, MA: Cascadilla Proceedings Project.

Miyamoto, Edson T. 2002. Case marker as clause boundary inducers in Japanese. *Journal of Psycholinguistic Research* 31(4). 307–347.

Miyamoto, Edson T. and Michiko Nakamura. 2003. Subject/object asymmetries in the processing of relative clauses in Japanese. In Gina Garding and Mimu Tsujimura (eds.), *Proceedings of the 22nd West Coast Conferences on Formal Linguistics*, 342–355. Somerville, MA: Cascadilla Press.

Miyamoto, Edson T. and Michiko Nakamura. 2005. Unscrambling some misconceptions: A comment on Koizumi and Tamaoka (2004). *Gengo Kenkyu* 128. 113–129.

Murasugi, Keiko. 2000. An Antisymmetry analysis of Japanese relative clauses. In Artemis Alexiadou, Paul Law, Andre Meinunger and Chris Wilder (eds.), *The syntax of relative clauses*, 231–263. Amsterdam: John Benjamins.

Nakamura, Michiko. 2003. *Processing of multiple filler-gap dependencies in Japanese*. Manoa, HI: University of Hawai’i dissertation.

Nakamura, Michiko and Edson T. Miyamoto. 2013. The object before subject bias and the processing of double-gap relative clauses in Japanese. *Language and Cognitive Processes* 28(3). 303–334.

Nakayama, Mineharu, Shravan Vasishth and Richard Lewis. 2006. Difficulty of certain sentence constructions in comprehension. In Mineharu Nakayama, Reiko Mazuka, and Yasuhiro Shirai (eds.), *Handbook of East Asian psycholinguistics 2: Japanese*. 277–284. Cambridge, UK: Cambridge University Press.

O’Grady, William. 1997. *Syntactic development.* Chicago, IL: University of Chicago Press.

Reali, Florencia and Morten H. Christiansen. 2007. Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language* 57(1). 1–23.

Roland, Douglas, Gail Mauner, Carolyn O'Meara and Hongoak Yun. 2012. Discourse expectations and relative clause processing. *Journal of Memory and Language* 66(3). 479–508.

Sakamoto, Tsutomu. 2015. Processing of syntactic and semantic information in the human brain: evidence from ERP studies in Japanese. In Mineharu Nakayama (ed.), *Handbook of Japanese Psycholinguistics*. Boston: De Gruyter Mouton.

Sakamoto, Tsutomu and Daichi Yasunaga. 2009. Ni kaku dōsi o fukumu kankeisetsu ni okeru shori-fuka o zōdai saseru gen'in nitsuite [The processing load of the relative clause including the dative verb]. *IECIE Technical Report* 109(140). 27–32.

Sato, Atsushi. 2011. *Nihongo kankeisetsu no shorifuka o kettei suru yōin no kentō: Kōpasu ni okeru shiyōhindo no eikyō o chūshin ni* [A study on determining factors for processing load of Japanese relative clauses: Influence of corpus-based frequency]. Hiroshima, Japan: Hiroshima University dissertation.

Sato, Atsushi, Barış Kahraman, Hajime Ono and Hiromu Sakai. 2009. Expectation driven by case markers: Its effect on Japanese relative clause processing. In Yukio Otsu (ed.), *The Proceedings of Tenth Tokyo Conference on Psycholinguistics*. 215–237. Tokyo: Hituzi Syobo.

Sato, Atsushi, Barış Kahraman and Hiromu Sakai. 2010a. Does frequency of occurrence make relative clause processing easier in Japanese? Poster Presented at *23rd* Annual CUNY Conference on Human Sentence Processing, New York University, New York, 16–20 March. [The poster can be downloaded from https://sites.google.com/site/kahramanbpapers]

Sato, Atsushi, Barış Kahraman and Hiromu Sakai. 2010b. Danwa kinō kara mita Nihongo-kankeisetsushori: Kōpasu chōsa to dokubunkeisokujikken ni yoru kenshō [Processing of Japanese relative clauses from viewpoint of discourse function: A study on the corpus analysis and self-paced reading experiment], *Proceedings of the 140th Meeting of Linguistic Society of Japan*. 212–217.

Sato, Atsushi, Barış Kahraman and Hiromu Sakai. 2012. When is the object relative clause easier to process than the subject relative clause? *Technical Report of IECIE* 112(145). 41–46.

Sawasaki, Koichi and Akiko Kashiwagi-Wood. 2015. Issues in L2 Japanese sentence processing: similarities/differences with L1 and individual differences in working memory. In Mineharu Nakayama (ed.), *Handbook of Japanese Psycholinguistics*. Boston: De Gruyter Mouton.

Schfriers, Herbert., Angela D. Friederici and Katja Kühn. 1995. The processing of locally ambiguous relative clauses in German. *Journal of Memory and Languag*e 34(4). 499–520.

Shibatani, Masayoshi. 1972. Semantics of Japanese causativization. *Foundations of Language* 9(3). 327–373.

Staub, Adrian. 2010. Eye movements and processing difficulty in object relative clauses. *Cognition* 116(1). 71–86.

Traxler, Matthew J., Rihana S. Williams, Shelley A. Blozis and Robin K. Morris. 2005. Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Languag*e 53(2). 204–224.

Ueno, Mieko and Suzan Garnsey. 2008. An ERP study of subject and object relative clauses in Japanese. *Language and Cognitive Process* 23(5). 646–688.

Wells, Justine B., Morten H. Christiansen, David S. Race, Daniel J. Acheson and Maryellen C. MacDonald. 2009. Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology* 58(2). 250–271.

Yamashita, Hiroko. 1997. The effect of word-order and case marking information on the processing of Japanese. *Journal of Psycholinguistic Research* 26(2). 163–188.

Yamashita, Hiroko, Laurie Stowe and Mineharu Nakayama. 1993. Processing of Japanese relative constructions. In Patricia Clancy (ed.), *Japanese/Korean linguistics* Vol. 2. 248–264. SLA, Stanford University. Distributed by Cambridge University Press.

Tsutomu Sakamoto

# 15 Processing of syntactic and semantic information in the human brain: Evidence from ERP studies in Japanese

## 1 Introduction

When a word is given in the course of an utterance, we comprehend the phonological, syntactic, and semantic properties of the word, locate the word into the structure thus far constructed, expect what comes next, combine all information together, and finally reach understanding of the whole utterance with consideration of pragmatic contexts. All these processes seem to be performed almost at once. Something amazing must be going on in our brain. Psycholinguistic researchers have been interested in these complicated language comprehension processes. What kind of information sources do we use in order to perform this amazing feat? When do we employ such information in the course of language processing? How do we handle such information for the purpose of language understanding? This chapter is devoted to answering these three questions (WHAT, WHEN, and HOW).

Recently, owing to both technological developments and theoretical advances, it has become possible to investigate the precise mechanism of these processes through the use of Event-Related brain Potentials (ERPs), which reflect electro-physiological activity in the brain in response to stimuli. ERP studies have shown that distinct brain responses appear to map onto distinct processes in response to linguistic stimuli. Research employing ERPs as an experimental tool for language processing (especially sentence comprehension) began in 1980 when Kutas and Hillyard published their pioneering study. In Japan, the first ERP studies on language comprehension appeared around 2000, when Hagiwara and her colleagues published a series of experimental papers (Hagiwara et al. 2000; Nakagome et al. 2001; Takazawa et al. 2002), and therefore the field of Japanese ERP research on language processing is still in its infancy. Since the amount of accumulated results is impoverished, there is still a lot to learn about the ERP indices of language processing in Japanese. This chapter aims to answer the above three questions by discussing possible dissociations and interactions of various linguistic processes in the human brain. We will review some ERP studies on syntactic and semantic processing in Japanese, and where possible, compare these studies to similar ones conducted in Indo-European languages such as English and German. Because of space limitation, we do not observe phonological information sources that are also very important (cf. Joo'o 1996; Koso, Ojima and Hagiwara 2011; Tamaoka et al. 2014).

The overall organization of this chapter is summarized as follows. In Section 2, we briefly introduce the methodology of ERPs. Section 3 examines the physiological evidence for dissociating syntactic and semantic processes as indexed by ERPs. In section 4 we investigate two information sources that are considered to affect the semantic processing. They are "semantic plausibility" (estimated by a "cloze probability") in a sentence and "semantic relatedness" (typically represented by a "category membership") between words. Section 5 illustrates that three sources affect the syntactic processing: "ungrammaticality", "garden-path", and "long-distance dependency". In Section 6 we reconsider the syntax-semantic dissociations and review two parsing models ("syntax-first" and "interactive") that attempt to account for the various types of linguistic processes. We will then conclude in Section 7 that the processing mechanism works in a "expectancy-driven" way.

## 2 What are Event-Related brain Potentials (ERPs)?

Before beginning our discussion of ERP studies of language processing, it is first necessary to briefly describe the ERP methodology. The human brain is made up of about a hundred billion neurons. When large groups of similarly aligned neurons fire, they generate small local electrical field potentials that are measurable at the scalp, which can be measured non-invasively. More precisely, we measure the changes in voltage over time, by placing electrodes on the surface of the scalp (see Figure 1).

The recorded changes in voltage over time, which are shown between +5 μV and -5 μV in this figure, are referred to as the electroencephalogram (EEG). The EEG reflects brain activity in response to the linguistic stimuli but also reflects processes

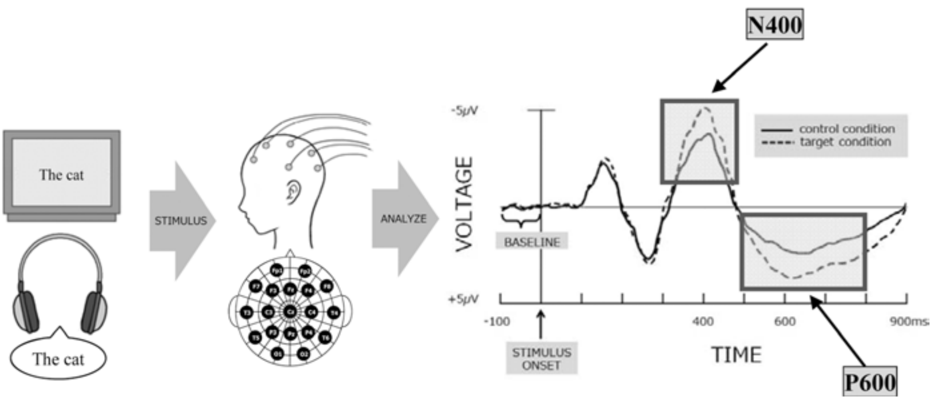
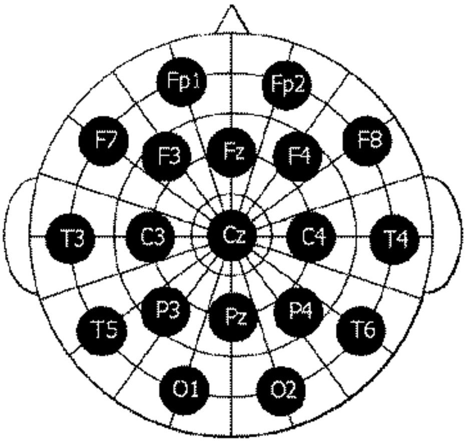



**Figure 1:** The flow from the presentation of visual and/or auditory stimulus to the ERP wave-forms. Details regarding the scalp distribution of electrodes are explained in Figure 2.




(i) Names of electrode positions

| | |
|---|---|
| Frontal pole | Fp1, Fp2 |
| Frontal | Fz, F3, F4, F7, F8 |
| Central | Cz, C3, C4 |
| Parietal | Pz, P3, P4 |
| Occipital | O1, O2 |
| Temporal | T3, T4, T5, T6 |

(ii) Names of electrode groups

| | |
|---|---|
| Midline | Fz, Cz, Pz |
| Anterior | Fp1, Fp2, Fz, F3, F4, F7, F8 |
| Posterior | Pz, P3, P4, O1, O2, T5, T6 |
| Parasagittal | F3, F4, C3, C4, P3, P4 |

**Figure 2:** The position of each electrode following the international 10–20 system. Note that this “10–20” system is not the “only” way of positioning the electrodes. There are some variations among laboratories.

completely unrelated to the experimental task. However, the purpose of neurolinguistic work is to isolate the brain’s response to the stimulus of interest by eliminating superfluous activity. Unfortunately, the changes in voltage in response to the stimuli of interest are very small, on the order of a few microvolts (μV). Therefore, these responses are usually masked by activity unrelated to the process being investigated (this unrelated activity is referred to as background EEG). In order to acquire EEG reflecting a specific mental process, it is necessary to record multiple responses that were time-locked to the onset of the stimulus of interest, and subsequently average these responses. The averaging process cancels out the majority of unrelated activity, and thus isolates the brain’s response to the stimulus in question. These extracted potentials are ERPs.

By convention, positive voltage is plotted down, and negative is plotted up.[1] The average voltage of 100 milliseconds (ms) prior to the onset of the critical stimulus is usually used as the “baseline” against which to examine the brain’s response to the stimuli. ERP responses are a multi-dimensional dependent measure. These dimensions include: (i) polarity (positive vs. negative), (ii) latency (time from stimulus onset), (iii) amplitude (voltage change) and (iv) distribution (on scalp). The positions for electrodes can be described by an international standard, shown in Figure 2.

The measurement of ERP effects is conducted in the following fashion. During a pre-specified time range (e.g., 300–500 ms), an experimental condition is compared

**1** Thus, there is no specific connotation in the term “negative” and “positive”. It is just an indication of the polarity of the waveforms.

to a control condition. For example, if the former elicits a negative response with higher amplitude than the latter, the target condition is said to have elicited enhanced negativity. This enhancement is assumed to index some cognitive process. The underlying assumption is that an increase in amplitude indexes an increase in some form of difficulty. The response with negative polarity peaking at around 400 ms after the onset of the target stimulus is called the "N400" effect and assumed to reflect the processing of semantic information (see Kutas and Federmeier 2011 for review). The response with positive polarity peaking at around 600 ms after the onset is called the "P600" effect and considered to index syntactic processing (see Bornkessel-Schlesewsky and Schlesewsky 2009 for review). Distinct from both the N400 and the P600 is the left anterior negativity or LAN, which appears to be closely tied to morpho-syntactic processing (see Molinaro, Barber and Carreiras 2011 for review) and to working memory (King and Kutas 1995; Kluender and Kutas 1993a, 1993b). Although the time window (300–500 ms) of the N400 and the LAN overlaps, the region of distribution distinguishes between the two components. Distinct from the broad distribution (mostly right posterior) of the N400, the distribution of LAN is restricted to left anterior regions of scalp (e.g., Fp1, F3, F7). In response to certain types of stimuli, biphasic N400–P600 or LAN-P600 patterns have also been observed (see Section 5.1.2). Although these classical functional interpretations have been much debated (as will be seen in subsequent sections), they serve here as a useful starting point for our discussion.

Since ERPs can be measured throughout the presentation of experimental stimuli, they provide a continuous, on-line record of the brain's electrical activity during language processing with very high temporal resolution. Furthermore, there is no need for participants to perform a secondary task, allowing researchers to investigate the brain's response during reading/listening in a more naturalistic setting.[2] Investigations of cognitive processes such as those involved in language processing should ideally be based on data that accurately reflects the process as it occurs in real time rather than with a dependent measure that is collected after the process has been completed. Thus, the ideal tool for investigating how language processing is developed over time would be on-line, continuous, and non-invasive; all properties of the ERP technique. However, the ERP method does have limitations in that it has relatively low spatial resolution. That is, it is difficult to specify the neuro-anatomical generator of the relevant ERP component. In order to improve the spatial resolution, we need neuroimaging methods such as functional Magnetic Resonance Imaging (fMRI) and Positron Emission Tomography (PET) (see Hagiwara's chapter in

2 While the EEG is recorded, participants can be asked to perform a secondary task such as lexical decision and grammatical judgment. This does not affect the ongoing ERP "recording" itself because the task is performed after the ERP recording has been completed. However, the task may or may not affect the underlying cognitive processes in question.

this volume on fMRI). Next, let us examine the syntactic and semantic processes indexed by these ERP components.

# 3 Syntax-semantics dissociations and ERPs

Many linguists operate under the basic assumption that syntactic phenomena are separable from semantic (and pragmatic) phenomena. Illustrating this independence, Noam Chomsky's famous example (1965: 149) "*Colorless green ideas sleep furiously.*" is a grammatical sentence of English, but semantically anomalous. This dissociation between syntax and semantics is widely accepted and implemented in many current linguistic theories. Using ERPs, we can observe electro-physiological evidence for this dissociation.

## 3.1 Syntax-semantics dissociations in English

Osterhout and Inoue (2007) showed that syntactic and semantic anomalies elicit distinct brain responses while participants are reading sentences like (1).[3] (Syntactic anomaly is marked with "*" and semantic anomaly with "?", and critical words are in boldface throughout this chapter.)

(1) a. Non-anomalous control sentences
*The cat will **eat** the food I leave on the porch.*
b. Semantically anomalous sentences
?*The cat will **bake** the food I leave on the porch.*
c. Syntactically anomalous sentences
**The cat will **eating** the food I leave on the porch.*

The ERP responses elicited by these three types of sentences are shown in Figure 3 below[4]. The anomaly in (1b) is semantic in the sense that *cats* cannot *bake* things, due to the non-human subject and the "selectional restrictions" the verb *bake* exerts on its subject. As is shown by the dashed line in Figure 3a, the semantic anomaly in sentence (1b) elicited a larger negative-going wave that peaked at about 400 ms (i.e., the N400 effect). On the other hand, the anomaly in (1c) is caused by the ungrammatical combination of the auxiliary *will* and the progressive form of the verb *eating*. The syntactic anomaly, as shown by the dashed line in Figure 3b,

3 In what follows, all participants are native speakers of the relevant language. When irrelevant to discussion, the details of the experiment such as procedures, statistics, participants' number, age, sex, etc. are omitted. Please consult the original papers for those and other details.

4 These figures are originally from Osterhout and Nicol (1999) and are slightly modified by Osterhout and Inoue (2007) for ease of explanation.

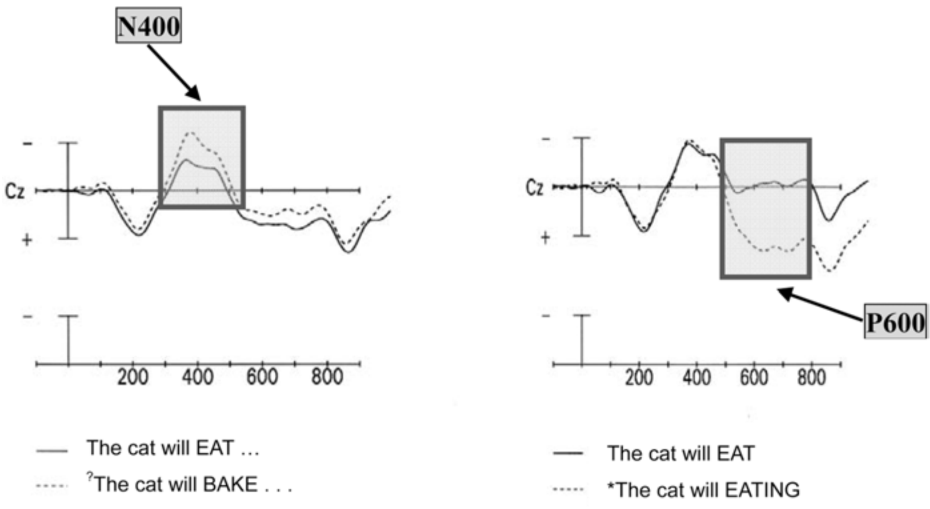



Figure 3a (semantic anomaly) Figure 3b (syntactic anomaly)

**Figure 3:** Grand-average ERPs recorded at Cz position (see Figure 2 for the location of Cz). Negativity is plotted up. Reproduced from Osterhout and Inoue (2007: 296) with permission of the publisher.

elicited a large positive wave with an onset at about 500 ms and a duration of several hundred ms (i.e., the P600 effect). Thus, this study (and many others as will be seen in subsequent sections) demonstrates that the dissociation between syntax and semantics is represented in the brain.

The examples shown here are a recent replication of some of the earliest findings in the ERP literature. Researchers are currently focusing on more complex constructions and manipulations in many languages. This endeavor has generated many new and interesting findings but also findings that are difficult to interpret and integrate with the existing literature. It is just this type of work that we discuss in the remainder of this chapter, focusing mainly on studies in Japanese.

## 3.2 Syntax-semantics dissociations in Japanese

The first issue is whether the type of syntax-semantics dissociation above is also observed in Japanese. We introduce two ERP experiments whose results serve as a base for subsequent discussions in this chapter.

### 3.2.1 Selectional restriction violations and wh-question violations

Nakagome et al. (2001) examined ERP responses elicited by the following two (semantic and syntactic) types of comparison.

(2) Semantic anomaly comparisons

a. Non-anomalous control sentences

| | | | | |
|---|---|---|---|---|
| *Taroo* | *ga* | ***ryokoo*** | *ni* | ***dekake-ta*.** |
| Taro | NOM | journey | DAT | set.out-PST |

'Taro set out on a journey.'

b. Semantically anomalous sentences (selectional restriction violation)

| | | | | |
|---|---|---|---|---|
| ?*Taroo* | *ga* | ***zisyo*** | *ni* | ***dekake-ta*.** |
| Taro | NOM | dictionary | DAT | set.out-PST |

'?Taro set out on a dictionary.'

(3) Syntactic anomaly comparisons

a. Non-anomalous control sentences

| | | | | |
|---|---|---|---|---|
| *Doobutuen* | *de* | ***nani*** | *o* | *mi-ta-***ka.** |
| zoo | LOC | what | ACC | see-PST-interrogative |

'What did you see in the zoo?'

b. Syntactically anomalous sentences (wh-question violation)

| | | | | |
|---|---|---|---|---|
| **Doobutuen* | *de* | ***nani*** | *o* | *mi-ta-***yo.** |
| zoo | LOC | what | ACC | see-PST-confirmative |

'*I saw what in the zoo.'

In (2a), the semantic relationship between the dative noun (*ryokoo* 'journey') and the verb (*dekake-ta* 'set out') is acceptable, whereas in (2b), the dative noun (*zisyo* 'dictionary') does not satisfy the selectional restriction imposed by the verb, and as such represents a semantic violation of the dependency between object and verb. In (3a), the dependency between the wh-phrase (*nani* 'what') and the sentence final question particle (Q-particle) "*-ka*" is syntactic rather than semantic, whereas in the syntactically anomalous (i.e., ungrammatical) sentence (3b) the wh-phrase is not followed by the requisite sentence-final Q-particle "*-ka*", but by a confirmative particle "*-yo*".

In the comparison between (2a) and (2b), the ERPs in the semantically anomalous condition elicited enhanced negativity (see Figure 4a) compared to the control condition during the latency range of the 300–700 ms (peak latency of 576 ms) at posterior sites as well as right temporal sites. Nakagome et al. interpreted this negativity as an N400, though the peak latency is later than is typically observed. On the other hand, the syntactically anomalous sentence (3b) elicited more positivity than the grammatical control sentence (3a) in the latency range of the 400–700 ms (peak latency of 616 ms) over relatively large areas of the scalp but most prominently at posterior sites. This positivity is similar to a canonical P600 effect (see Figures 3b and 4b). Thus, they concluded that the semantic anomaly elicited an N400 effect and the syntactic anomaly a P600 effect. These results suggest that the brain is

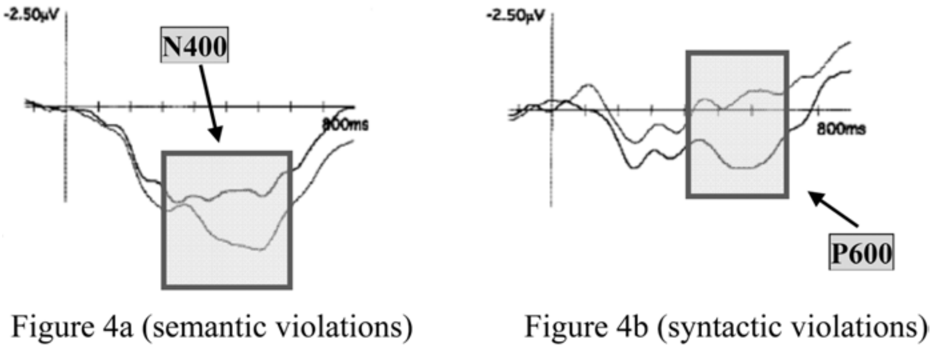



**Figure 4:** Grand-averaged ERPs in the anomalous and control conditions. Thick lines indicate the anomalous conditions, and thin lines the control conditions. The electrode of Figure 4a is located between T4 and T6. The electrode of Figure 4b is located between Pz and C4. Reproduced from Nakagome et al. (2001: 307) with permission of the publisher and authors.

indeed responding differently to semantic violations (indexed by the N400) compared to syntactic ones (indexed by the P600), identical to the pattern we observed in the English examples (1) above.

### 3.2.2 Selectional restriction violations and case-assignment violations

In the research mentioned above, it should be noted that the non-anomalous control sentences (2a) and (3a) for the semantic and syntactic anomalies are two very different types of constructions. This property of the materials makes it difficult to directly compare the two violation types. To control for this confound, it is necessary to construct an experimental design similar to (1) in English. In pursuit of this aim, Nashiwa, Nakao and Miyatani (2007) examined ERP responses to sentences including semantic and syntactic anomalies, crucially comparing these anomalies to the same control sentences (cf. Friederici, Steinhauer and Frisch 1999; Osterhout and Nicol 1999). Examples are provided below.

(4) a. Non-anomalous control sentences
*Taroo ga* ***ringo*** *o* ***tabe-ta***.
Taro NOM apple ACC eat-PST
'Taro ate an apple.'

b. Semantically anomalous sentences (selectional restriction violation)
?*Taroo ga* ***reizooko*** *o* ***tabe-ta***.
Taro NOM refrigerator ACC eat-PST
'?Taro ate a refrigerator.'

c. Syntactically anomalous sentences (case-assignment violation)

| | | | | |
|---|---|---|---|---|
| **Taroo* | *ga* | *ringo* | ***ni*** | ***tabe-ta***. |
| Taro | NOM | apple | DAT | eat-PST |

'*Taro ate to an apple.'

The semantic manipulation is the same as in Nakagome et al. (2001). That is, (4b) violates the selectional restriction between the verb *tabe-ta* 'ate' and its object *reizooko* 'refrigerator' because the lexical-semantic property of the verb *eat* requires its object to be edible, and a refrigerator is usually not edible in our real world. Compared to the control sentence (4a), the semantic anomaly in (4b) elicited N400 effects (see Figure 5). Note that the syntactic anomaly in (4c) differs from the anomaly of ungrammatical wh-questions in Nakagome et al.'s (3b). In Japanese, the transitive verb *tabe-ta* 'ate' requires an object that is marked with an accusative case marker "*-o*", rather than the dative case marker "*-ni*" that appears in the anomalous sentences. We will refer to this pattern as a "case-assignment violation" between verbs and objects. Nashiwa, Nakao and Miyatani (2007) observed a P600 effect in response to this syntactic (and/or morpho-syntactic) violation in (4c) compared to the grammatical sentence (4a) as in Figure 5. As such, they have provided evidence for a syntax-semantics dissociation in Japanese with an experimental design parallel to that in Osterhout and Inoue (2007).

This section shows the syntax-semantics dissociation in Japanese is virtually identical to that in English. The three studies reviewed elicited N400 effects in response to selectional restriction violations in English and Japanese. In the next section, we will examine the nature of N400 effects more extensively, and discuss recent theoretical interpretations of this effect.

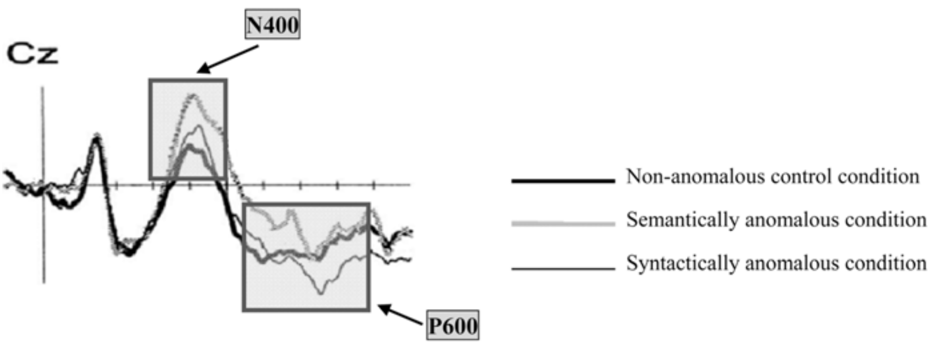



**Figure 5:** The ERPs of the final verbs for three different types of sentence at Cz. A line irrelevant to our discussion is omitted from the original figure. Reproduced from Nashiwa, Nakao and Miyatani (2007: 315) with permission of the publisher and authors.

# 4 Semantic processing indexed by the N400

The N400 appears only when there is a semantic violation (anomaly, incongruity) such as selectional restriction violations. However, this simplification does not accurately represent the processes that the N400 indexes. There are various information sources that elicit N400 effects (see Kutas and Federmeier 2011 for review). Among these, we discuss two important sources in this section: a "plausibility effect" in sentential context that is estimated by a measure called a "cloze probability", and a "semantic relatedness", which is typically represented by a "category membership" between two lexical items in isolation.

## 4.1 Semantic plausibility effects on semantic processing: cloze probability

In many studies it has been demonstrated that N400 amplitude is inversely related to the ease of accessing lexical information for the forthcoming word that is expected from the preceding context: the less expected the word is, the larger the N400 effect will be. Ease of access is estimated by a measure referred to as "cloze probability" that is calculated with a pen-and-paper task in which participants are asked to complete the missing final word of a sentence: the more participants choose a particular word, the higher the cloze probability that word is said to have.

Kutas, Lindamood and Hillyard (1984) examined how this cloze probability affects the amplitudes of the N400 using three types of sentence as follows.

(5) a. Best (most probable) completion sentences
*She called her husband at his* ***office***.

b. Less probable completion sentences
$^{?}$*Captain Sheir wanted to stay with the sinking* ***raft***.

c. Semantically anomalous sentences
$^{??}$*George was fired but he couldn't tell his* ***fog***.

The best (i.e., "most probable") completion sentences such as (5a) ended with the most probable word expected from the preceding sentential context. The less probable completion sentences such as (5b) are meaningful but relatively improbable. Their cloze probability was smaller than that of the best completion sentences. Semantically anomalous sentences such as (5c) ended up with a nonsense message. The starting point of the ERP recording was time-locked at the beginning of the presentation of each final word. The results are shown below.

The N400 amplitudes to semantically anomalous words in (5c) were larger than those elicited by the best (most probable) completion words in (5a) and the less

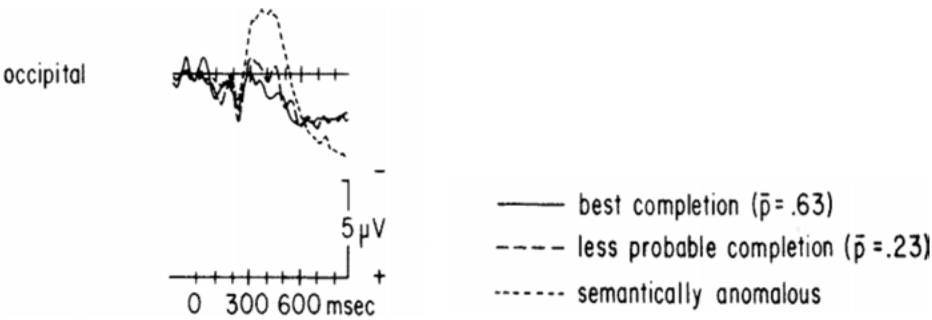



**Figure 6:** The grand average ERPs at the occipital (O1, O2) position elicited by the best, less probable, and anomalous words at the end of each sentence. Reproduced from Kutas, Lindamood and Hillyard (1984: 226) with permission of the publisher.

probable completion words in (5b). The less probable completions yielded a N400 amplitude in between the other two sentences. The N400 amplitude is sensitive to cloze probability. Thus, Kutas, Lindamood and Hillyard (1984) claim that there is "a systematic decline in the N400 amplitudes for terminal words as a function of increasing cloze probability" (p. 233).

It is evident that semantic violation is not a necessary condition for N400 elicitation even if it can be a sufficient condition. The N400 amplitudes increase as an inverse function of cloze probability, which measures the relative degree of expectancy for upcoming words by native speakers. Therefore, the N400 can be considered to reflect the mental process of expectation in semantic processing, indicating that the language comprehension mechanism works in an "expectancy-driven" way.

Now, reconsider the Japanese examples in Section 3.2. The sentential contexts (*Taroo ga ryokoo/zisyo ni* and *Taroo ga ringo/reizooko o*) make us strongly expect the upcoming (and not yet presented) words in the examples. In the non-anomalous control sentences the encountered final verbs are expected, whereas in the anomalous sentences the sentence final verbs are not expected ones resulting in implausible sentences. Considering the previous study in English by Kutas, Lindamood and Hillyard (1984), we can argue that this decrease of semantic plausibility in Japanese examples elicited N400 effects. That is, what causes the elicitation of the N400 effects is not the semantic (selectional) violation but the low plausibility due to unsatisfied expectation in sentential contexts.

## 4.2 Semantic relatedness effects on semantic processing

When the expectation was not satisfied by the crucial word, N400 effects are modulated according to the degree of cloze probability. In other words, the more plausible the sentence becomes, the less amplitude of the N400 is elicited. However, there are some cases where sentential contexts do not affect the modulation of N400 effects.

### 4.2.1 Sentential contexts irrelevant to N400 modulations: Truth value

Fischler et al. (1983) found that a "true" negative sentence like *A sparrow is not a vehicle* elicited larger N400 than a "false" negative sentence such as ?*A sparrow is not a bird*. That is, in this case it was not the false but the true statement that elicited the N400 effect. Because the semantically anomalous statement failed to elicit the N400, semantic anomaly is not sufficient for yielding an N400 activity. This finding, together with the discussion in Section 4.1, may suggest that semantic anomalies are neither necessary nor sufficient for N400 elicitation.

Fischler et al. (1983) claim that this pattern of results is due to the detection of the semantic mismatch between *sparrow* and *vehicle* being performed earlier than the judgment of the truth value of the sentence as a whole. Note that there are (at least) three distinct processes involved: (i) processing the affirmative-negative distinction, (ii) processing the semantic relatedness between the two nouns, and (iii) verifying the truth value of the sentence as a whole. An important point is that both the processing of semantic relatedness and the truth/false judgment occur at the sentence final position. Thus, the ERP component observed at this position could be reflecting two superimposed processes and should therefore be interpreted with caution.

In order to disentangle the processes of calculating semantic relatedness and verifying the truth value of sentences, Katayama, Miyata and Yagi (1987) investigated Japanese sentences because the predicates are sentence-final. They performed an ERP experiment employing the same design as Fischler et al. (1983) except that the word order in Japanese is "$NP_1$–$NP_2$–verb" instead of "$NP_1$–verb–$NP_2$" in English, enabling the authors to potentially investigate the processes in question independently as they occur at different sentence positions in their materials.

(6) a. True affirmative: non-anomalous sentences
***Suzume*** *wa* ***tori*** *dearu.*
sparrow TOP bird is
'A sparrow is a bird.'

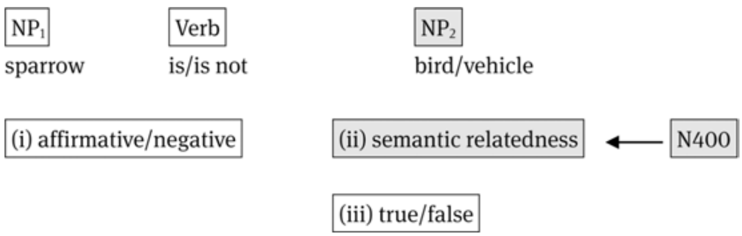
b. False affirmative: semantically anomalous sentences
?***Suzume*** *wa* ***norimono*** *dearu.*
sparrow TOP vehicle is
'A sparrow is a vehicle.'

c. True negative: non-anomalous sentences
***Suzume*** *wa* ***norimono*** *denai.*
sparrow TOP vehicle is.not
'A sparrow is not a vehicle.'

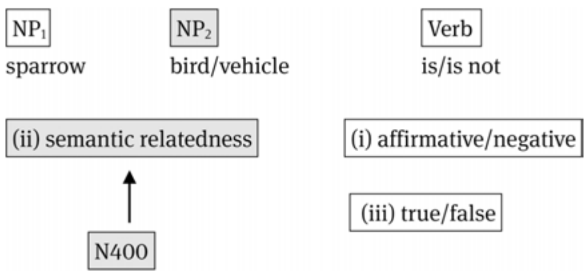
d. False negative: semantically anomalous sentences
?***Suzume*** *wa* ***tori*** *denai.*
sparrow TOP bird is.not
'A sparrow is not a bird.'

At the position of the second noun, the incongruent "*suzume – norimono*" ('sparrow – vehicle') combinations (6b) and (6c) elicited a larger N400 response compared to the congruent "*suzume – tori*" ('sparrow – bird') combinations (6a) and (6d). The predicate did not yet appear at this sentence position, so readers could not tell whether the sentence as a whole was affirmative or negative and therefore could not judge whether the statement was true or false. Thus, this N400 effect solely reflects the processing difficulty caused by the semantic incongruity between *suzume* 'sparrow' and *norimono* 'vehicle'. This result duplicates the finding by Fischler et al. (1983) while eliminating the possible confounding factor described above. Note that at the sentence final position in Japanese, however, the two processes of affirmative/negative discrimination and true/false judgment are confounded. In the experimental design of Fischler et al. (1983) with English stimuli, both the processing of the semantic relatedness and that of true/false judgment occur at sentence final position so that the two distinct processes are necessarily confounded. The cases in English and Japanese can be summarized as follows.

(7) a. Truth value verification sentences in English




b. Truth value verification sentences in Japanese




The fact that Japanese is a head-final language made these findings possible. It shows the importance of studying multiple languages in order to make conclusions

about the mapping between ERP responses and underlying cognitive processes. Only when combining results of head-initial and head-final languages can we dissociate two confounding processes. And then, we can claim that N400 effects are not elicited by truth value or sentence meaning, but the relatedness between the first and the second noun. This is a good example of a cross linguistic study from which we can learn a lot about ongoing language processing in the human brain.

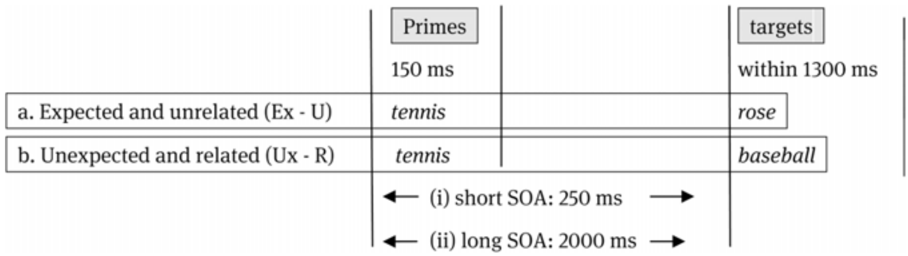
### 4.2.2 Priming effects on the N400: automatic versus controlled process

In Section 4.1, we have examined N400 effects in which sentential contexts generate expectancies for upcoming linguistic elements. As shown in the true/false verification paradigm, however, the incongruent combination of two words (e.g., *sparrow* – *vehicle*) elicits larger N400 effects than the congruent one (e.g., *sparrow* – *bird*). In other words, the sentential frame is not the necessary condition for eliciting N400 effects. These effects on the N400 are basically equivalent to lexical priming effects in the sense that the preceding elements in these sentences (essentially "primes") facilitate or inhibit the access to subsequent elements (essentially "targets").

Holcomb (1988) found that the unrelated word pairs (e.g., *table* – *animal*) elicited significantly larger N400 amplitude than the related word pairs (e.g., *nurse* – *doctor*) in the situation where automatic process is assumed. Thus, Holcomb claims that attentional process is not a necessary ingredient for eliciting this priming N400 effect (see also Kutas and Hillyard 1989). In Holcomb's experimental design, however, the prime words were overtly presented. Since the properties of primes had already been consciously processed before participants performed the lexical decision task, the semantic information of the primes must have influenced the controlled process. In order to eliminate this possible influence, Brown and Hagoort (1993) presented primes with masked and unmasked conditions. Masked primes are not consciously perceived by participants so that controlled processes cannot play a role. The N400 effects were observed only in the unmasked condition but not in the masked condition, which indicates that the priming N400 effect is not observed if conscious identification of the prime is not possible. Thus they claim that the N400 reflects the controlled lexical integration process (see also Chwilla, Brown and Hagoort 1995).

It is claimed that there are two underlying mechanisms responsible for controlled processes: expectancy-induced priming and post-lexical integration (Neely 1991). Nakao and Miyatani (2007) investigated the separate contribution of these two mechanisms. Their experimental design can be summarized as follows.

(8) Lexical priming experiment in Nakao and Miyatani (2007) with translated Japanese words in English




Participants were required to decide whether the target word is a correct Japanese word or nonword. Targets were presented until participants responded with a key press within a 1300 ms limit. There were two different types of stimuli set: "expected and unrelated (Ex-U)" and "unexpected and related (Ux-R)" (cf. Neely 1977). In the Ex-U condition, when a prime belonging to a certain category was presented (e.g., *tenisu* 'tennis' belonging to a sport category), participants were instructed to anticipate receiving a target from a different category (e.g., *bara* 'rose' belonging to flower category). This instruction was intended to evoke the expectancy effect among participants. They were forced to expect an incoming word that is unrelated to the preceding word. In the Ux-R condition, a target of the same category member as the prime was presented (e.g., *yakyuu* 'baseball'). This target is semantically related to the prime, but it was not the expected target due to the instruction that induces the participants to anticipate a target belonging to a different category from the prime. In this experimental design, the expectancy effects and semantic relatedness effects seem to be successfully separated. There were two different conditions concerning stimulus onset asynchrony (SOA): short SOA (250 ms) and long SOA (2000 ms). The SOA represents the duration between the onset of the prime and that of the target.

The results showed that both expectancy effects and semantic relatedness effects were indexed by the N400 effects in both the short and long SOA conditions. The ASA (automatic spreading activation) may account for the semantic relatedness effect in the short SOA condition, but it cannot explain this effect in the long SOA condition because the ASA is assumed to occur as an early automatic process that should have already disappeared in the long SOA condition. Although the expectancy-induced priming can account for the expectancy effect in the long SOA condition, it cannot explain this effect in the short SOA condition because the time range of this short condition is not long enough to generate the expectancy set, that is a time-consuming process. Thus, the authors claim that the modulation of the N400 effect in both of the two SOA conditions can only be explained by the view that the N400 indexes the post-lexical integration process.

Some researchers claim that the N400 primarily indexes underlying processes in which we may automatically (i.e., without conscious awareness) access the semantic information of a word. Others, on the other hand, argue that the N400 reflects the controlled process where the critical word is matched with the preceding word and is integrated into the higher conceptual constituent. Since there is ample experimental evidence that supports both of these claims, it seems to be difficult to maintain the view that the N400 reflects either an automatic or a controlled process. Therefore, Kutas and Federmeier (2011) claim that the N400 does not reflect a single process but a complex of (at least) two distinct processes and propose a hybrid theory to explain what processes the N400 indexes. According to Kutas and Federmeier, the wave form of the unrelated (unexpected) word starts to diverge from that of the related (expected) word around 200 ms from the onset of critical word. They claim that the process performed from this diverging point to the peak latency of the N400 (i.e., around 400 ms) reflects an automatic process, and the process from the peak to the point where the two wave forms converge indexes a controlled process. Although a more detailed investigation is necessary to prove the validity of this claim, the fine-grained sensitivity of the ERP measure could make it possible in the future.

### 4.2.3 N400 effects in a phrase level

When a prime and a target belong to different categories, the semantic relatedness becomes feeble. One important factor to determine the relatedness is whether the prime and the target belong to the same category, although there are various types of relatedness such as synonymy, antinomy, etc. Category mismatch between primes and targets enlarges the amplitude of the N400 compared to the cases without such a mismatch. The relationship between the two words is crucial to the modulation of N400 effects. In priming experiments, however, little attention is paid to whether the two words constitute one linguistic constituent or not. In the priming experiments examined above, both primes and targets are nouns (e.g., *tennis* – *baseball*), and they are just juxtaposed. The two nouns are not integrated to constitute one element. On the other hand, for example, a sequence of two nouns can constitute one complex word such as *door-knob*, *house-keeper*, etc. Similarly, a noun preceded by an adjective can constitute one phrase like *red color*. Here, we examine two studies related to N400 priming effects on phrase level constituents using category mismatch in Japanese. The first study concerns category mismatches between different sensory expressions and the second one examines category mismatches between classifiers and their attendant nouns.

#### 4.2.3.1 Sensory mismatches in Japanese

The first study concerning N400 effects in a phrase level of Japanese is Sakamoto et al. (2003). It found N400 effects in response to mismatches between a modifying

adjective and an NP in sensory expressions (In the following examples, double question mark "??" indicates more severe anomaly than the single question mark "?").

(9) a. Congruent visual expression (visual adjective + visual noun)
*akai iro*
red color

b. Incongruent visual expression (tactile adjective + visual noun)
?*namerakana iro*
smooth color

c. Congruent tactile expression (tactile adjective + tactile noun)
*namerakana tezawari*
smooth touch

d. Incongruent tactile expression (visual adjective + tactile noun)
??*akai tezawari*
red touch

Combinations of words related to different sense modalities produce mismatch expressions like (9b) ?*namerakana iro* 'smooth color' and (9d) ??*akai tezawari* 'red touch' in contrast to normal expressions like (9a) *akai iro* 'red color' and (9c) *namerakana tezawari* 'smooth touch'. Sakamoto et al. (2003) conducted a five-scale questionnaire test for comprehensibility (acceptability) judgments whose results showed that (i) mismatch expressions (9b) and (9d) are less comprehensible than normal expressions (9a) and (9c), and that (ii) there are two patterns between mismatch expressions: the "high-comprehensible mismatch expression (HighCME)" such as (9b) ?*smooth color* and the "low-comprehensible mismatch expression (LowCME)" like (9d) ??*red touch*. The first finding seems to be plausible because these mismatch expressions involve category mismatches between two different sensory domains. The second finding is consistent with observations that the direction of combination affects the comprehensibility of sensory mismatch expressions in English (Ullmann, 1951; Williams, 1976) and in Japanese (Kusumi, 1988; Sakamoto, 1983). That is, an expression is easier to understand when the modifying direction is "upward" (from a lower modality such as tactile to a higher modality such as visual) than when it is "downward" (from higher to lower modalities). Thus, some upward expressions such as *soft color*, *cool color*, and *warm color* are very familiar idioms (or, dead metaphors) in both English and Japanese (and perhaps in many other languages).

In order to verify these two findings obtained from the judgment task, Sakamoto et al. (2003) conducted ERP experiments. An adjective (e.g., *red*) was presented for 700 ms. After a 600 ms interval, a noun (e.g., *color*) was presented for 600 ms. After the presentation of the second word, participants were asked to judge the comprehensibility of each phrase on a scale of 1–5. The results of this post-hoc judgment

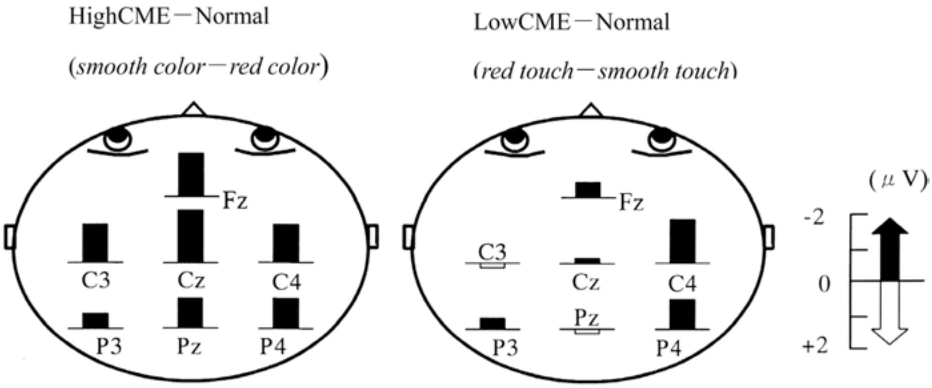



**Figure 7:** Hemispheric dissociation of N400 amplitudes in sensory mismatch expressions (High/LowCME minus Normal expressions). Negativity is indicated by a black bar. Reproduced from Sakamoto et al. (2003: 391) with permission of the publisher and authors.

task were the same as the previous questionnaire test mentioned above. The results of ERP experiments revealed that (i) the mismatch expressions produced a significantly larger N400 amplitude than the normal expressions, but that (ii) the difference between High/LowCMEs was not reflected in the amplitudes of the N400.

The authors suggest that the topographical difference of the N400 could be related to the difference between High/LowCMEs: HighCMEs such as ?*smooth color* increased N400 amplitudes in the whole region, while LowCMEs like ??*red touch* produced activation only in the right hemisphere as shown in Figure 7.

It is well known that the right and left hemispheres have different functions. Brownell et al. (1984) pointed out that patients with right-hemisphere damage were insensitive to connotative meanings (e.g., *cold* has a connotative meaning of *remoteness*). See also Anaki, Faust and Kravetz (1998) and Winner and Gardner (1977) for the importance of the right hemisphere to understand metaphorical expressions. In understanding very familiar and highly conventional metaphors like *broken heart*, however, it is reported that the left hemisphere is also at work (Giora 1999). Thus, it might be the case that a LowCME like ??*red touch* is processed in the right hemisphere as an unfamiliar metaphor, while a HighCME like ?*smooth color* is processed in the left and right hemisphere as a familiar metaphor. The topographical difference of the N400 may reflect the dissociation of underlying processes between HighCMEs and LowCMEs. It thus appears that the N400 is a useful measure to investigate the hemispheric difference of language processing in general.

#### 4.2.3.2 Classifier mismatches in Japanese

In a similar vein, but employing a category mismatch specific to Japanese, Sakai et al. (2006) observed N400 effects in response to a mismatch between nouns and

classifiers. In Japanese when we count an entity (e.g., *inu* 'dog') a specific classifier (e.g., *-biki*) with numerals is used as shown in the following examples (CLF= Classifier).

(10) a. Congruent noun-classifier construction
***inu*** *san-**biki***
dog three-CLF
'three (animal entity of) dogs'

b. Incongruent noun-classifier construction
??***inu*** *san-**bon***
dog three-CLF
'??three (long-slender entity of) dogs'

When expressing something in Japanese that is equivalent to *three dogs* in English, the expression would be *inu san-biki* 'dog three animal entity' or, *san-biki no inu* 'three animal entity of dog' in Japanese. That is, the correct construction in the Japanese counting system requires a specific classifier to be attached to numerals. We cannot simply say **san inu* 'three dogs' or **inu san* 'dogs three' in Japanese. Compared to the congruent expression *inu san-biki*, the incongruent phrase ??*inu san-bon* elicited an enhanced N400 effect. After encountering the noun *inu* 'dog', it is difficult to access and/or to integrate the classifier *-bon* that modifies long-slender entities such as pencils, sticks, etc. (e.g., *enpitu san-bon* 'three long-slender entity of pencil'). The increase of N400 amplitude presumably indexes increased difficulty in the access and/or integration processes between numeral classifiers and corresponding nouns.

In classifier languages such as Japanese (Chinese, Korean, etc.), classifiers overtly denote which noun belongs to which category. For instance, *dog* belongs to the animal category while *pencil* belongs to the category of long-slender entities. In contrast, languages such as English (German, French, etc.) do not overtly indicate whether *dog* and *pencil* belong to different categories. Even in languages without classifier systems, however, native speakers do know that a *dog* belongs to the animal category including *cat*, *monkey*, etc. and is not a category member of *pencil*, *stick*, etc. Independent of overt linguistic apparatus to indicate these relationships, there seems to be some universal conceptual structure among human beings. The sensitivity of the N400 to this classifier-type category mismatch can shed light on the conceptual architecture of the mental lexicon in general.

## 4.3 Summary of Section 4

We have mentioned two information sources that modulate N400 effects: semantic plausibility and semantic relatedness. First, we observed that the amplitude of the

N400 is inversely correlated with semantic plausibility, which is estimated by a cloze probability that reflects the degree of expectancy induced by sentential contexts. The stronger the expectancy for an incoming item is, the smaller the amplitude of the N400 response becomes. It is true that semantic violations (anomaly, incongruity) can modulate the amplitude of the N400, but the violation itself is neither a necessary nor a sufficient condition for elicitation of N400 effects. The observations described here point to a more subtle state of affairs in which the degree to which an incoming element is plausible (expected, predicted) modulates the amplitude of the N400 in a more graded and subtle pattern. As such, modulations of N400 amplitude reflect brain activity associated with graded semantic plausibility correlated with expectancy.

Second, we mentioned that semantic relatedness (or lack thereof) modulates N400 amplitudes. As shown in the truth-value experiments, N400 modulations can be affected not by sentence meaning or truth value but the semantic relatedness (association) between the first and the second noun. Even a single word activates semantically related words in the mental lexicon. The relatedness is typically represented by a category membership, although there are other types of semantic relatedness. Both sentential and single word contexts facilitate the expectation of upcoming words. When we encounter these expected words, the amplitudes of N400 are reduced compared to unexpected and less-expected words.

What we have learned from the ERP studies thus far can be summarized as follows. We use semantic information sources such as semantic plausibility (estimated by cloze probability) and semantic relatedness (typically represented by a category membership) during the time window of around 400 ms in an expectancy-driven way. This is the first tentative answer to the question (WHAT, WHEN, and HOW) that we posited in the Introduction. However, we concentrated on only two (but important) instances of information sources that modulate N400 effects. There are various other linguistic sources such as frequency, word class, pseudo words, etc. and also non-linguistic factors such as faces, pictures, gestures, etc. See Kutas and Federmeier 2011 for review.

# 5 Syntactic processing indexed by the P600 (and related ERP components)

This section observes three types of syntactic processing difficulty in Japanese and some other languages. Section 5.1 deals with "ungrammaticality" that requires "syntactic repair": incorrect wh-questions and case-assignment violations. Section 5.2 concerns processing difficulty caused by "garden-path" sentences as an illustration of "syntactic revision". As an illustration of "syntactic integration", in Section 5.3, we examine "long-distance dependency" between two distant elements in a sentence.

## 5.1 Syntactic repair processes required by ungrammaticality

In section 3, we introduced two instances of syntactic anomalies (i.e., ungrammaticality). One is the ungrammatical ending of wh-questions in Nakagome et al. (2001) and the other is object-verb case-assignment violation in Nashiwa, Nakao, and Miyatani, (2007). First, we examine the former type of ungrammaticality, and then we move on to the latter type of case-assignment violation more extensively.

### 5.1.1 Ungrammatical wh-questions in Japanese

Now, let us return to the wh-questions investigated in Nakagome et al. (2001) mentioned at Section 3.2.1. The relevant examples are reproduced below.

(3) a. Non-anomalous control sentences

| | | | | |
|---|---|---|---|---|
| *Doobutuen* | *de* | ***nani*** | *o* | *mi-ta-**ka**.* |
| zoo | LOC | what | ACC | see-PST-interrogative |

‘What did you see in the zoo? ’

b. Syntactically anomalous wh-questions

| | | | | |
|---|---|---|---|---|
| **Doobutuen* | *de* | ***nani*** | *o* | *mi-ta-**yo**.* |
| zoo | LOC | what | ACC | see-PST-confirmative |

‘*I saw what in the zoo. ’

The wh-element (*nani* ‘what’) requires the interrogative Q-particle (*-ka* or *-no*). In the ungrammatical sentence (3b), this requirement is not satisfied as the confirmative particle *-yo* appears in its place, leading to the elicitation of P600 effects, presumably because a long-distance dependency could not be formed due to the lack of the requisite second element.

When the expected particle is not encountered, the result is an ungrammatical sentence. Even in this situation, we may try to rescue the sentence by repairing this ungrammaticality and salvaging a coherent interpretation. For example, we may attempt to replace the affirmative particle *-yo* to the interrogative particle *-ka* (or *-no*), or re-interpret the wh-word *nani* ‘what’ to a non-wh homophone *nani* ‘something special’ (differing from the wh-word in pitch accent). The cognitive processes involved in this type of repair are unknown, but based on the literature this is at least a plausible hypothesis. The crucial point in this argument is that the P600 indexes the mental process of repairing the syntactic violation. Then, let us call this type of P600 a “repair P600” (cf. Kaan and Swaab 2003). Since there are presumably multiple types of P600 effects that index various cognitive processes as will be discussed in subsequent sections, henceforth we will label P600 effects based on the underlying processes that they are assumed to index. Note that the terms here

are employed for expository convenience and may not be widely adopted within the literature.

### 5.1.2 Case-assignment violation processes indexed by ERPs

In the previous subsection, we observed that the P600 is assumed to reflect the repair process to rescue the ungrammaticality caused by syntactic violations between wh-elements and corresponding question particles. In the following subsections we examine some studies concerning case-assignment violations in Japanese.

#### 5.1.2.1 Case-assignment violation processes in English and German

Before going into the case-assignment violation issues in Japanese, however, let us briefly consider the case-assignment violations in English and German.

(11) Case-assignment violations in English (Coulson, King and Kutas 1998)

a. Grammatical sentences (transitive verb + accusative pronoun)
*The plane took* ***us*** *to paradise and back.*

b. Ungrammatical sentences (transitive verb + nominative pronoun)
**The plane took* ***we*** *to paradise and back.*

(12) Case-assignment violations in German (Friederici and Frisch 2000)

a. Grammatical accusative sentences (accusative object + accusative verb)

| | | | | | | |
|---|---|---|---|---|---|---|
| *Anna* | *weiß,* | *dass* | *der* | *Kommissar* | ***den*** | *Banker* |
| Anna | knows | that | $\text{the}_{[\text{NOM}]}$ | inspector | $\text{the}_{[\text{ACC}]}$ | banker |

| | | |
|---|---|---|
| ***abhorte*** | *und* | *wegging.* |
| $\text{monitored}_{[\text{ACC}]}$ | and | left. |

'Anna knows that the inspector monitored the banker and left.'

b. Ungrammatical dative sentences (accusative object + dative verb)

| | | | | | | |
|---|---|---|---|---|---|---|
| **Anna* | *weiß,* | *dass* | *der* | *Kommissar* | ***den*** | *Banker* |
| Anna | knows | that | $\text{the}_{[\text{NOM}]}$ | inspector | $\text{the}_{[\text{ACC}]}$ | banker |

| | | |
|---|---|---|
| ***beistand*** | *und* | *wegging.* |
| $\text{helped}_{[\text{DAT}]}$ | and | left. |

'*Anna knows that the inspector helped the banker and left.'

The ungrammatical English sentence (11b) with violations in pronominal case marking, compared to grammatical sentence (11a), elicited both the LAN and P600 (this is called a "biphasic pattern"). The ungrammatical German sentences such as (12b) elicited a LAN followed by a P600 effect compared to the grammatical sentences

like (12a) at the matrix verb position. These results are very suggestive in that the case-assignment violation elicits both LAN and P600 effects. That is, the processing of case-assignment violations may involve two distinct processes: morphological checking process that is reflected by the LAN and syntactic repairing process that is indexed by the P600. The results of English and German are consistent, which may imply the universality of case violation processing.

However, note that in these experiments the comparison is between the accusative pronoun (*us*) and the nominative pronoun (*we*) in English and the accusative verb (*abhorte* 'monitored') and the dative verb (*beistand* 'helped'). Since the comparison was performed between the two different types of nouns and verbs in question, the ERP responses associated with these comparison may reflect differences of lexical types in addition to the processing of case-assignment violations. Furthermore the typological similarity between English and German may have induced a similar pattern of language processing. Now, let us go into the case-assignment violation issues in Japanese. The properties of head finality and overt case marking in Japanese are different from English and similar to German, and Japanese is typologically different from both English and German.

#### 5.1.2.2 Case-assignment violation processes indexed by the P600

The type of case-assignment violations in Nashiwa, Nakao and Miyatani (2007) is reproduced below.

(4) a. Non-anomalous control sentences

| | | | | |
|---|---|---|---|---|
| *Taroo* | *ga* | *ringo* | ***o*** | ***tabe-ta***. |
| Taro | NOM | apple | ACC | eat-PST |

'Taro ate an apple.'

c. Syntactically anomalous sentences

| | | | | |
|---|---|---|---|---|
| **Taroo* | *ga* | *ringo* | ***ni*** | ***tabe-ta***. |
| Taro | NOM | apple | DAT | eat-PST |

'*Taro ate to an apple.'

The sequence of "NP *ga* NP *o*" (nominative – accusative) generates the expectation that the upcoming verb can subcategorize its argument as an accusative. The verb *tabe-ta* 'ate' satisfies this expectation as in (4a). When we are given the sequence of "NP *ga* NP *ni*" (nominative – dative), we expect to encounter a verb such as *au* 'meet' that can take a dative object (e.g., *Taro ga Hanako ni atta*. 'Taro met Hanako.'). That is, a dative object generates an expectation for a dative verb. This expectation is not satisfied by an accusative verb such as *tabe-ta* 'ate' in (4c). This case-assignment violation between a dative object and an accusative verb elicited P600 effects without preceding LAN effects. Thus, the LAN effects observed

in English (*us/*we*) and German (*abhote/*beistand*) could be caused by the comparison between two different types of lexical item. Another possibility, however, is that the experimental design in Nashiwa, Nakao and Miyatani (2007) caused the lack of LAN effects. Remember that the experiment was designed to compare syntactic and semantic violations. Since the semantic anomaly (selectional restriction violation) was so salient that the LAN effect in the case-assignment violation, even if it does exist, may have been undetectable. In what follows, we explore these two possibilities for the lack of LAN effects in Japanese, manipulation the experimental sentences and designs.

#### 5.1.2.3 Case-assignment violation processes indexed by the N400 and P600

There are two types of object case markers (accusative "***-o***" and dative "***-ni***") and two types of transitive verb (accusative verb and dative verb). Manipulating these two types of case markers and transitive verbs, Kobayashi et al. (2007) examined the following four types of sentences. This 2 × 2 experimental design could perhaps more clearly reveal the ERP responses associated with the process of object-verb case mismatches observed in the previous studies.

(13) a. Grammatical accusative sentences (accusative object + accusative verb)

| | | | | | | |
|---|---|---|---|---|---|---|
| *Hookago* | *sensei* | *ga* | *seito* | ***o*** | *kyoositu de* | ***sikatta***. |
| after.school | teacher | NOM | student | ACC | classroom-at | scolded |

'After school, the teacher scolded the student at the classroom.'

b. Ungrammatical accusative sentences (dative object + accusative verb)

| | | | | | | |
|---|---|---|---|---|---|---|
| **Hookago* | *sensei* | *ga* | *seito* | ***ni*** | *kyoositu de* | ***sikatta***. |
| after.school | teacher | NOM | student | DAT | classroom-at | scolded |

(14) a. Grammatical dative sentences (dative object + dative verb)

| | | | | | | |
|---|---|---|---|---|---|---|
| *Basu-de* | *zyookyaku* | *ga* | *suri* | ***ni*** | *suguni* | ***kizuita***. |
| bus-at | passenger | NOM | pickpocket | DAT | immediately | noticed |

'At the bus, the passenger noticed the pickpocket immediately.'

b. Ungrammatical dative sentences (accusative object + dative verb)

| | | | | | | |
|---|---|---|---|---|---|---|
| **Basu-de* | *zyookyaku* | *ga* | *suri* | ***o*** | *suguni* | ***kizuita***. |
| bus-at | passenger | NOM | pickpocket | ACC | immediately | noticed |

In this experimental design, participants are able to process the case-assignment violation without possible effect from semantic violation. Furthermore, it is possible to compare the grammatical and ungrammatical sentences at the same lexical item: the sentence final verb. The authors report that the ungrammatical sentences in both accusative and dative sentences elicited a biphasic pattern consisting of an N400

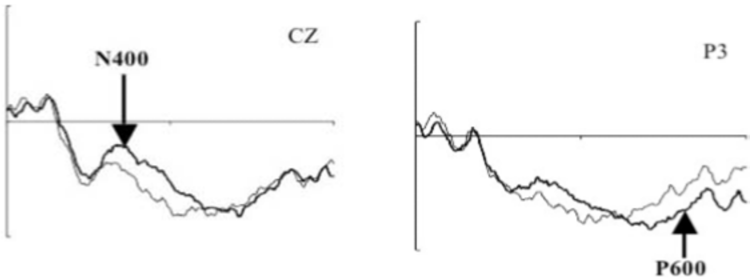



**Figure 8:** Grand Average ERPs of Cz and P3. Thin lines indicate the grammatical sentences (accusative object + accusative verb), and thick lines the ungrammatical sentences (dative object + accusative verb). Negativity is plotted up. Reproduced from Kobayashi et al. (2007: 48) with permission of the publisher and authors.

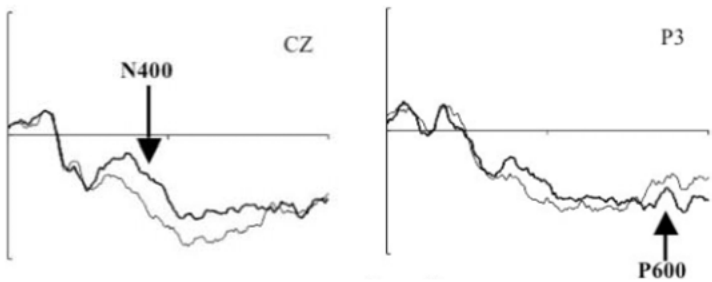



**Figure 9:** Grand Average ERPs of Cz and P3. Thin lines indicate the grammatical sentences (dative object + dative verb), and thick lines ungrammatical sentences (accusative object + dative verb). Negativity is plotted up. Reproduced from Kobayashi et al. (2007: 47) with permission of the publisher and authors.

followed by a P600 at the sentence final verb. See Figure 8 for (13) and Figure 9 for (14). They interpreted the N400 as an index of violation checking, a process that requires lexical access in order to determine whether the verb in question is a dative or accusative verb. Thus, they claim that the same process is involved in the checking of case-assignment violations, independent of the particular case involved. And they interpreted the subsequent P600 effect as indexing repair operations after a violation has been checked.

Since the experimental design in Kobayashi et al. (2007) allows us to examine the processing of syntactic case violation eliminating the possible semantic violation effects and comparing at the same lexical item, it may be plausible to claim that the case-assignment violation in Japanese is processed in a two-step fashion as in English and German. Thus, it seems to be plausible to conclude that the experimental design to compare syntactic and semantic violations in Nashiwa, Nakao and Miyatani (2007) caused the lack of LAN effects. The salient semantic violation may have

obscured the detection of the negativity effects in the morpho-syntactic violation. This explanation is further supported by the English example (1) by Osterhout and Inoue (2007) which was designed to examine the difference between syntactic and semantic violations as in Nashiwa, Nakao, and Miyatani (2007). The morpho-syntactic violation such as **The cat **will eating** the food* elicited the P600 effect without detectable LAN/N400 effects.

Considering the results of English and German, this morpho-syntactic violation is expected to elicit the LAN. Very interestingly, however, this violation did not modulate LAN effects but N400 effects. As a reason for the N400 effects, Kobayashi et al. mentioned that there is a tight relationship between grammatical cases and thematic roles. Nominative case-marked NPs tend to bear the role of Agent, accusative-marked NPs tend to bear the role of Patient, and dative NPs the role of Goal. In (13b), for instance, the verb *sikatta* 'scolded' requires its object to be a Theme with an accusative case. Contrary to this requirement, the object actually bears a Goal role with a dative case. This clash between the Theme and the Goal is assumed to generate an anomaly at the level of thematic-role assignment, and therefore the authors claim that we would expect to observe the ERP correlates of thematic anomaly, i.e., the N400 (Friederici and Frisch 2000; Frisch and Schlesewsky 2001, 2005).

Since the case-assignment violation in both English and German did not elicit the N400 but the LAN, the N400 effects in this type of case-assignment violation could be specific to Japanese. Furthermore, Arao et al. (2007) reported that the "dative object + accusative verb" elicited a LAN effect, while the "accusative object + dative verb" in contrast elicited an N400. The results may therefore imply that the violation between a dative object and accusative verb is morpho-syntactic as in English and German cases, while the violation between the accusative object and dative verb is instead related to semantic/thematic processes as is claimed by Kobayashi et al. (2007). It is an empirical question requiring further investigation as to whether the differences in ERP effects reflect the distinct processes between accusative and dative constructions during the process of case-assignment violations.

#### 5.1.2.4 Summary of case-assignment violation processes

The studies reviewed here provide evidence that the human brain is sensitive to, and responds to, case-marking inconsistencies between verbs and the arguments that they select for. All studies except Nashiwa, Nakao and Miyatani (2007) exhibited a biphasic pattern of ERP responses. A possible explanation for the lack of LAN/N400 effects could be that the experimental design prevented negativity effects from being detected because of the salient semantic violation. Since Kobayashi et al. (2007) and Arao et al. (2007), which eliminated the possible influence of a semantic violation, exhibited the biphasic ERP patterns, it is probable that the case-assignment violation in Japanese is processed by two steps as in English and German: violation checking at the morphological level and repair of the violation at the syntactic level.

Unfortunately, however, no clear-cut pattern emerges across studies, and we therefore cannot definitively conclude what processes these components index. In order to resolve these apparent inconsistencies, it will be necessary in the future to conduct more studies of this type with carefully controlled materials that can incrementally eliminate potential accounts of the cognitive processes involved.[5]

### 5.1.3 Summary of syntactic repair processes

We observed that illegitimate affirmative particles instead of interrogative particles result in ungrammatical wh-questions in Section 5.1.1. The long-distant relationship between wh-phases and Q-particles was not completed. This type of long-distance ungrammaticality elicited a P600 effect that is called the "repair P600", which indicates that we try to make an incoherent message coherent by repairing the ungrammaticality. Section 5.1.2 was devoted to discussions on ungrammaticality caused by morpho-syntactic case-assignment violations. This is a local dependency between verbs and objects. This type of local ungrammaticality elicited the repair P600 as well as a preceding negativity (N400/LAN). This biphasic pattern of ERP responses indicate that the morpho-syntactic violation is processed in a two-step fashion: checking of the violation at the morphological level and repair of the violation at syntactic level. ERP studies gave us physiological evidence that ungrammaticality is an instance of information source in syntactic processing. However, it has been suggested in the literature that it is not the ungrammaticality but the processing cost that is the main source for eliciting the P600 effects. In Sections 5.2 and 5.3, we observe such processing difficulty without ungrammaticality.

## 5.2 Syntactic revision processes required by garden-path

In the ERP literature on syntactic processing, it has been mentioned that not only ungrammatical sentences but also sentences with high syntactic processing costs

---

**5** In this subsection we discussed case-assignment violations between the accusative/dative objects and accusative/dative verbs. Different from this type of violation, Mueller, Hirotani, and Friederici (2007) investigated sentences in which case-assignment violations were triggered by illicit sequences of case-marked nouns. For example, they investigated sentences in which two NPs bore nominative case markers (*kamo* ***ga*** *nezumi* ***ga*** 'duck NOM mouse NOM') or two accusative case markers (*kamo* ***o*** *nezumi* ***o*** 'duck ACC mouse ACC'). The results of the experiment showed that a sequence of NPs with the same case marker elicited a biphasic N400-P600 pattern at the second NP. Frisch and Schlesewsky (2001, 2005) also reported a similar pattern of responses to this type of serial case-marking violation. However, Díaz et al. (2011) reported only a P600 effect in response to a sequence of two ergative marked NPs in Basque, a pattern that was also reported by Bise and Sakamoto (2011) in response to a sequence of two accusative-marked NPs in Japanese. Though reviewing the literature in its entirety is beyond the scope of this chapter, the issue of this type of case-assignment violation is an interesting and open question.

(such as the "garden-path" sentences discussed below) elicit the P600 effect. In this section, we will consider this type of P600 effect, which we term the "revision P600".

### 5.2.1 Revision processes in English

As stated above, not only ungrammatical sentences but also grammatical but non-preferred syntactic constructions are known to elicit P600 effects. For example, Osterhout, Holcomb and Swinney (1994) conducted an experiment comparing the following four types of sentences.

(15) a. Grammatical control sentences (intransitive verb)
*The doctor **hoped** the patient was lying.*

b. Ungrammatical sentences (transitive verb)
**The doctor **forced** the patient was lying.*

c. Grammatical and preferred sentences (intransitively biased verb)
*The doctor **believed** the patient was lying.*

d. Grammatical but non-preferred sentences (transitively biased verb)
*The doctor **charged** the patient was lying.*

The intransitive verb *hope* can take only a complement clause, while the transitive verb *force* can take only an NP as its argument. In the ungrammatical sentence (15b), the ungrammaticality becomes clear at the point of the auxiliary *was* where the P600 effect over the posterior site of scalp was observed compared to its grammatical counterpart (15a). The observed P600 effect is presumably the same as the "repair P600" discussed in Section 5.1. In order to obtain a coherent interpretation, we have to repair the ungrammatical structure that had been built.

On the other hand, (15c) and (15d) are both grammatical because the verb *believe* and *charge* can take both complement clauses and NPs. That is, these verbs are ambiguous in terms of (in)transitivity. However, *believe* tends to be used as an intransitive verb with a complement clause and *charge* is often used as a transitive verb with an object NP. While (15c) is more frequent and preferred, (15d) is less frequent and therefore poses more of a processing challenge. If ungrammaticality is the only source of the P600 effect, these two types of grammatical sentences should not be different concerning the modulation of the P600 effect. Contrary to this prediction, at the position of the auxiliary *was*, the grammatical but non-preferred sentence (15d) elicited a larger P600 effect than the grammatical and preferred sentence (15c). The authors argue that in (15d) we would initially attach *the patient* as the object of the transitively biased verb *charged*. Upon encountering the auxiliary *was*, we are forced to abandon this structure and build a new one in which the

second NP becomes the subject of a complement clause, a revision process known as the "garden-path". The authors claim that the garden-path effect occurred at (15d) because we immediately take into account the verb-specific information such as (in)transitivity.

The results showed that both ungrammatical and garden-path sentences elicited P600 effects. However, the distribution of the P600 between them is different. The ungrammatical sentences have posterior distribution and the non-preferred garden-path sentences more frontal distribution. Thus, Osterhout, Holcomb and Swinney (1994) revealed the distributional evidence that the P600 reflects the cognitive process involved in the revision process necessitated by garden-path sentences (see also Osterhout and Holcomb 1992, 1993). Here, we will refer to these types of effect as "revision P600s".

Hagoort, Brown and Osterhout (1999) argue that the revision P600 has a frontal distribution and the repair P600 has a posterior distribution and therefore that these two types of P600 effect reflect a different cognitive processes. Kaan and Swaab (2003), however, claim that there is no distributional difference between the revision and repair P600s: both have posterior distributions, and that the difference between these two types of effect is actually reflected in differences in amplitude. According to Kaan and Swaab (2003), the cognitive processes involved in repair and revision are essentially the same, although the cost of repair is higher than that of revision. Additional research is required to determine whether the repair and revision processes are qualitatively distinct or essentially the same but associated with the quantitative difference of processing costs.

### 5.2.2 Revision processes in Japanese

Osterhout, Holcomb and Swinney (1994) claim that we make immediate use of verb-specific information to construct a preferred sentence structure. The preferred interpretation of *the patient* in (15d) is the object of the transitively biased verb *charge*. The cause of the preference was the verb-specific information that this verb is transitively biased. Japanese sentences provide an interesting test of this claim because verbs are sentence final and therefore verbal information cannot be employed to construct the preferred structure. Oishi, Yasunaga and Sakamoto (2007) investigated the revision process in the following garden-path sentences in Japanese.

(16) a. Simple transitive sentences taking two arguments

| | | | | |
|---|---|---|---|---|
| [$_{NP}$ *Daizin* | ***ga***] | [$_{NP}$ *honbu* | *ni* | *atumatta* |
| minister | NOM | headquarters | LOC | donated |

| | | |
|---|---|---|
| *uragane* | ***o***] | ***nusunda***. |
| secret.money | ACC | stole |

'A minister stole the secret money donated to the headquarters.'

b. Ditransitive sentences taking three arguments

[$_{NP}$ *Daizin* ***ga***] [$_{NP}$ *honbu* ***ni***] [$_{NP}$ *atumatta*
minister NOM headquarters DAT donated

*uragane* ***o***] ***azuketa***.
secret.money ACC entrusted

'A minister entrusted the secret money donated to him to the headquarters.'

It is generally assumed that we build the simplest structure possible on the first-pass through the sentence processing (Frazier and Fodor 1978). According to this principle we would first construct a simple mono-clausal structure when the first verb *atumatta* is encountered, assuming that this verb is intransitive (and equivalent to the intransitive verb *gather* in English as shown in (17) below). Note that the sequence "*honbu ni*" has presumably been identified as a postpositional phrase (PP) and been attached as an adjunct of the matrix clause at this point in the parse.

(17) *Daizin* *ga* [$_{VP}$ [$_{PP}$ [$_{NP}$ *honbu*] *ni*] *atumatta*]
minister NOM headquarter LOC gathered
'Ministers gathered at the headquarter.'

Upon encountering the next word *uragane o* 'secret money ACC', however, we recognize that the initial simple mono-clausal analysis must be abandoned because it is impossible to incorporate this word into the structure that has thus far been built. It is clear that the simple sentence has to be revised to a complex relative clause structure, yet there are still two possible structures that we can revise to as shown in (18) below, where "ec" indicates an empty category and IP is an Inflectional Phrase, i.e., a sentence.

(18) a. *Daizin* ***ga*** [$_{VP}$ [$_{NP}$ [$_{IP}$ ec$_i$ [$_{VP}$ [$_{PP}$ [$_{NP}$ *honbu*] *ni*] *atumatta*]]
*uragane*$_i$ ***o***] ...(*nusunda*/**azuketa*) ]

b. *Daizin* ***ga*** [$_{VP}$ [$_{NP}$ *honbu* ***ni***] [$_{NP}$ [$_{IP}$ ec$_i$ [$_{VP}$ *atumatta*]]
*uragane*$_i$ ***o***] ... (**nusunda*/*azuketa*) ]

The difference between (18a) and (18b) is caused by the two different interpretations of the post nominal particle *ni*. Sadakane and Koizumi (1995) note that *ni* has two possible grammatical functions in Japanese: either as a postposition or a dative case marker. According to Miyagawa (1989), this difference is crucial because postpositions (such as *kara* 'from', *ni* 'at, on, to') project their own maximal projections in the syntax, while case markers (such as *ga* 'NOM', *ni* 'DAT') do not project but are rather incorporated into NPs as clitics.

(19) a. Postpositions

[$_{\mathrm{PP}}$ [$_{\mathrm{NP}}$ John] *kara*], [$_{\mathrm{PP}}$ [$_{\mathrm{NP}}$ John] *ni*]

John from, John at (on, to)

b. Case markers

[$_{\mathrm{NP}}$ John *ga*], [$_{\mathrm{NP}}$ John *ni*]

John NOM, John DAT

When the post-nominal particle *ni* is interpreted as a postposition, it possesses lexical meaning and therefore can assign a thematic role to its argument NP. Thus the sequence *honbu ni* 'at headquarters' forms a postpositional phrase as in (18a). If, on the other hand, *ni* is interpreted as a case marker that does not have semantic content, the particle *ni* does not take an argument and cannot assign a thematic role. Thus the sequence *honbu ni* 'headquarter DAT' in (18b) forms an NP as one constituent.

If we have no preference for one of these two possible constructions, there would be no difference in terms of processing difficulty. If we prefer (18a), a simple transitive verb like *nusunda* 'stole' will incur no additional processing cost because the two arguments (*daizin* ***ga*** and *uragane* ***o***) required by this transitive verb are present in the structure. When the ditransitive verb *azuketa* 'entrusted', which requires three arguments, is subsequently encountered, however, we diagnose that the indirect object of the verb (a dative-marked NP) is absent. Therefore, we are forced to reanalyze the sequence "*honbu ni*" from an adjunct analysis to a dative NP analysis so that the requirement of the ditransitive verb is satisfied. On the other hand, if we have a preference for the ditransitive structure as in (18b), which has three arguments (*daizin* ***ga***, *honbu* ***ni***, and *uragane* ***o***), the simple transitive verb *nusunda* 'stole' will incur additional processing cost because this verb cannot take three arguments. The ditransitive verb *azuketa* 'entrusted' does not engender such difficulty because of a straightforward match between the existing three arguments and the argument-taking properties of the ditransitive verb. Oishi, Yasunaga and Sakamoto (2007) conducted an ERP experiment to determine whether these two structures are processed differently, and if so, whether the revision process is triggered at the transitive or the ditransitive verb, localizing the position at which additional processing difficulty is incurred.

As shown in Figure 10, the results revealed that simple transitive verbs such as (16a) elicited a P600 effect compared to their ditransitive counterparts like (16b), which could plausibly be interpreted as indexing the difficulty associated with the revision process at (16a), indicating that this process is triggered by the transitive verb. The revision P600 was observed at sentence-final transitive verbs, suggesting that we have in fact already pursued a preferred ditransitive construction (NOM – DAT – ACC) before verb-specific information has become available. Since the absence of a verb until this position eliminates any source of verb information as to which structure should be built, the preference for the ditransitive construction (16b)

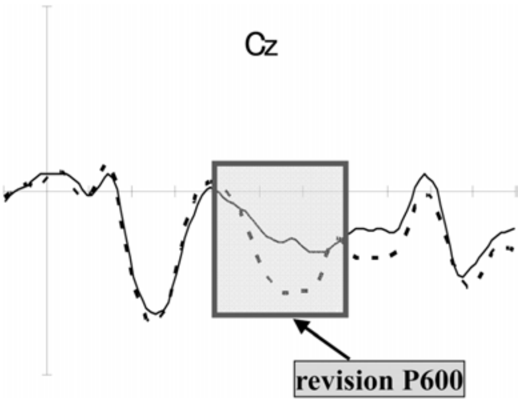



**Figure 10:** Grand Average ERPs at Cz elicited by the matrix verbs of the transitive and ditransitive verb sentences. ERPs to transitive verbs are plotted with dotted lines, and those to ditransitive verbs are plotted with solid lines. Negativity is plotted up. Reproduced from Oishi, Yasunaga, and Sakamoto (2007: 376) with permission of the publisher and authors.

over the simple transitive construction (16a) must be based on a purely structural measure as is discussed below. Previous studies claim that we will maintain as much existing structure as possible in the revision process (Frazier and Clifton 1998; Sturt and Crocker 1996). This structure-preserving principle would prefer (16a) and (18a) to (16b) and (18b) because the structural dependency between the adjunct and the verb ([$_{VP}$ [$_{PP}$ [$_{NP}$ *honbu*] *ni*] *atumatta*]), which has been already constructed as in (17), is preserved in (16a) and (18a) where "*honbu ni*" is interpreted as an adjunct PP. However, this parsing principle cannot account for the experimental results that the ditransitive sentences such as (16b) and (18b) are preferred over simple transitive sentences like (16a) and (18a). The authors, then, propose a principle called the "Error Signal-based Revisions Principle (ESRP)" in order to provide an account for the results. The ESRP is a simplicity metric which incorporates the head noun *uragane o* 'secret money ACC' into the existing structure. The relative head noun signals that the presumed simple sentence analysis is incorrect and therefore that this head noun must have been extracted from somewhere in the structure. Consider the structures where the relative heads are relocated into the original empty category (ec) positions. The structure [$_{IP}$ *Uragane ga* [$_{VP}$ *atumatta*]] in (16b) and (18b) is much simpler than the structure [$_{IP}$ *Uragane ga* [$_{VP}$ [$_{PP}$ [$_{NP}$ *honbu*] *ni*] *atumatta*]] in (16a) and (18a) because the former does not include the adjunct PP "*honbu ni*."

### 5.2.3 Summary of syntactic revision processes

So, it appears to be the case that language-specific properties dictate the strategies that we pursue when attempting to build structure from ambiguous input. As shown

in Osterhout, Holcomb and Swinney (1994), it appears to be the case that in a head-initial language like English, native speakers exploit verb-specific information in order to predict and pursue the construction of a preferred structure. However, the results by Oishi, Yasunaga and Sakamoto (2007) suggest that native speakers of head-final languages like Japanese appear to be pursuing a different strategy, employing simplicity metrics to determine which structure is in fact preferred before the verb-specific information becomes available. The difference of parsing strategies between the two languages became clear by investigating the head-initial and -final characteristics of the constructions in question. Remember the discussion in Section 4.2.1 where the head-final property of Japanese made it possible to eliminate the experimental confounding in truth value verification sentences. Again, we see another instance in which the high temporal resolution of the ERP technique has enabled us to investigate the underlying mechanism of language processing as syntactic structure is built in real time.

## 5.3 Syntactic integration processes required by long-distance dependency

The previous section discussed the P600 effects indexing the cost of syntactic revision from grammatical and preferred (expected) structures to grammatical but non-preferred (unexpected) ones. Since the revision P600 is observed in both grammatical sentences, ungrammaticality is not a necessary condition for the elicitation of P600 effects. In this section, we examine another case in which ungrammaticality is irrelevant to P600 effects. An element early in the sentence generates an expectation that is satisfied by a downstream element. The satisfaction of this expectation has also been observed to elicit P600 effects that may reflect the difficulty of the syntactic integration processes as will be seen in many studies mentioned in the following subsections.

### 5.3.1 Integration processes in English wh-constructions

Examining various types of wh-constructions, Kluender and Kutas (1993a, 1993b) argued that the maintenance and the retrieval of a filler (e.g., a moved wh-element) is indexed by the LAN (left anterior negativity) effect and Kaan et al. (2000) claimed that the integration process between a filler and a gap (e.g., an original position of a moved wh-element) is reflected by the P600 effect. Furthermore, Phillips, Kazanina and Abadac (2005) investigated how these ERP components are affected by manipulating the distance between a filler and a gap. This manipulation of dependency length helps to clarify which subprocesses these ERP components reflect. Consider the following examples (CP = Complementizer Phrase = clause).

(20) a. Short-distance that-clause sentences
*The detective hoped that the lieutenant knew* [$_{\text{CP}}$ ***that*** [$_{\text{IP}}$ *the shrewd witness would recognize the accomplice in the lineup*]].

b. Short-distance wh-clause sentences
*The detective hoped that the lieutenant knew* [$_{\text{CP}}$ ***which accomplice$_i$*** [$_{\text{IP}}$ *the shrewd witness would recognize* **<gap>$_{\mathbf{i}}$** *in the lineup*].

(21) a. Long-distance that-clause sentences
*The lieutenant knew* [$_{\text{CP}}$ ***that*** [$_{\text{IP}}$ *the detective hoped* [$_{\text{CP}}$ *that* [$_{\text{IP}}$ *the shrewd witness would recognize the accomplice in the lineup*]]]].

b. Long-distance wh-clause sentences
*The lieutenant knew* [$_{\text{CP}}$ ***which accomplice$_i$*** [$_{\text{IP}}$ *the detective hoped* [$_{\text{CP}}$ *that* [$_{\text{IP}}$ *the shrewd witness would recognize* **<gap>$_{\mathbf{i}}$** *in the lineup*]]].

In both the short- and long-distance sentences, an incomplete wh-dependency elicited a sustained anterior negativity (sAN), starting immediately after the wh-phrase until the embedded verb *recognize*. And also at this embedded verb position, a P600 effect was observed in the wh-clause sentences compared to the that-clause sentences. There was no difference in the amplitude of the P600 between short- and long-distance sentences, but the P600 had an earlier onset in the short-distance sentences appearing around 300–500 ms than in the long-distance sentences around 500–700 ms. Based on these results, the authors proposed two subprocesses of integration. One is the retrieval (reactivation) of the head of a syntactic dependency, which is length-sensitive. Since the filler has been held in working memory, it has to be retrieved for integration. Thus, the retrieval is a kind of pre-integration process. The longer syntactic dependency becomes the later P600 responses appear. The process of retrieving a filler is not reflected in the amplitude but in the onset of the response. After the retrieval process, another subprocess occurs in which a thematic-role assignment and a compositional semantic interpretation are performed. These truly integrative processes are claimed to be length-insensitive.

In the study of English wh-constructions by Phillips, Kazanina and Abadac (2005), it was shown that there are (at least) three subprocesses involved in the processing of filler-gap dependencies. (i) On encountering a wh-element, we recognize it as a filler and hold it in working memory until the gap position, which is indexed by the sAN. (ii) We retrieve the filler from working memory, which is length-sensitive and indexed by the early P600. (iii) At the verb position, a thematic-role assignment and compositional integration occur, which is length-insensitive and indexed by the late P600.

### 5.3.2 Integration processes in German wh-constructions

Wh-elements in English are basically ambiguous concerning their thematic roles (and cases). It has been claimed that this is why the filler must be integrated at the gap position where thematic roles are assigned. According to Phillips, Kazanina and Abadac (2005), the length-insensitive (and genuine) integration process is considered to consist of two subprocesses: the thematic-role assignment and the subsequent compositional process between the dislocated wh-phrase and the verb. It is not clear whether both of these subprocesses are necessary to elicit P600 effects or whether one of them is solely (or mostly) responsible for the P600 effect.

Contrary to English, wh-phrases in German and Japanese are case marked so that thematic roles and grammatical functions are fairly unambiguous. We can predict the thematic roles of wh-phrases before encountering the thematic-role assigning verb. If we posit a syntactic gap position for a dislocated wh-element as soon as possible, the integration would occur prior to a verb position. In order to test this prediction, let us examine German indirect wh-constructions examined by Fiebach, Schlesewsky and Friederici (2001).

(22) a. Subject wh-constructions

*Thomas fragt sich,* [$_{CP}$ ***wer$_i$*** [$_{IP}$ **<gap>$_i$** *am Mittwoch nachmittag*
Thomas asks himself who$_{[NOM]}$ on Wednesday afternoon

*nach dem Unfall* ***den*** *Doktor verständigt hat*]].
after the accident the$_{[ACC]}$ doctor called has

'Thomas asks himself who has called the doctor on Wednesday afternoon after the accident.'

b. Object wh-constructions

*Thomas fragt sich,* [$_{CP}$ ***wen$_i$*** [$_{IP}$ *am Mittwoch nachmittag*
Thomas asks himself who$_{[ACC]}$ on Wednesday afternoon

*nach dem Unfall* ***der*** *Doktor* **<gap>$_i$** *verständigt hat*]].
after the accident the$_{[NOM]}$ doctor called has

'Thomas asks himself who the doctor has called on Wednesday afternoon after the accident.'

In (22a), the wh-word *wer* 'who$_{[NOM]}$' is in nominative case and assumed to have moved to the clause initial interrogative position, leaving a trace (i.e., gap) in subject position of the embedded clause. In (22b), on the other hand, the wh-word *wen* 'who$_{[ACC]}$' is in accusative case, and is assumed to have moved to the clause initial interrogative position while leaving a trace (gap) in object position of the embedded clause. In (22a) since the wh-filler *wer* 'who$_{[NOM]}$' is adjacent to its gap and fills the gap immediately, the memory load should be very low. In (22b), on the other hand, the filler *wen* 'who$_{[ACC]}$' is distant from its gap. We must keep the filler in working

memory until an appropriate gap appears. The results showed that the object wh-construction (22b) elicited a sustained left anterior negativity (sLAN), in comparison to the subject wh-construction (22a), onsetting at the first prepositional phrase (*am Mittwoch* 'on Wednesday') and persisting until the subject noun phrase (*der Doktor* 'the$_{[NOM]}$ doctor'). This result is consistent with that of Kluender and Kutas (1993a, 1993b) and Phillips, Kazanina, and Abadac (2005). The authors concluded that the sLAN effect reflects the cost of storing the wh-filler in working memory until the syntactic dependency between the dislocated filler and its gap can be established.

Furthermore, the object wh-construction (22b) elicited a P600 effect at the second NP (*der Doktor*) position. Note that this effect was observed at the pre-gap NP position before the thematic-role assigning verb appears. When participants encountered the accusative marked wh-word *wen*, they may have started to search for an accusative gap at the earliest position. Since the "subject (nominative) – object (accusative) – verb" is the canonical word order in German embedded clauses, the first possible position for the object gap is immediately after the nominative case marked subject. That is, participants may have assumed the construction "object (accusative)$_i$ – subject (nominative) – <gap>$_i$ – verb". If this is the case, the P600 is considered to reflect the process of syntactic integration between the dislocated wh-element and syntactic gap before a thematic role is assigned by a verb. The accusative marked *wen* 'who$_{[ACC]}$' is more likely to have a Theme role rather than an Agent role. Since the morphological information of the wh-element provides enough information concerning thematic role (and case), we do not have to wait until the verb that actually assigns the thematic role (and case). As the lexical-semantic information of the verb is not available at the pre-gap NP position, the integration process should be purely syntactic. At the clause-final verb position, we would only check the semantic congruency between the wh-element and the verb. No P600 was elicited at the verb position because this semantic checking process occurs in both the subject and object wh-constructions. In order to test whether the integration P600 is really observed at the pre-gap NP position before the verb appears, let us consider the scrambling phenomena in Japanese that allows fairly free word order.

### 5.3.3 Integration processes in Japanese scrambled sentences

In a construction of long-distance dependency such as a scrambling construction, the interpretation of a scrambled element (i.e., a filler) depends on its original position (i.e., a gap). Hagiwara et al. (2007) compared two types of complex sentences in Japanese: one with canonical and the other with noncanonical word order (see also Koso, Hagiwara and Soshi 2007; Ueno and Kluender 2003). Consider the following examples that have the same propositional meaning with different word order.

(23) a. Canonical conditions (CC)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Kaiken* | *de* | $[_{CP}$ *syatyoo* | *wa* | $[_{CP}$ *hisyo* | *ga* | *bengosi* | *o* |
| meeting | at | president | TOP | secretary | NOM | lawyer | ACC |

| | | |
|---|---|---|
| *sagasiteiru* | *to*] | *itta*]. |
| look.for | COMP | said |

‘At the meeting, the president said that the secretary was looking for the lawyer.’

b. Scrambled conditions (SC)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Kaiken* | *de* | ***bengosi***$_i$ | ***o*** | $[_{CP}$ *syatyoo* | *wa* | $[_{CP}$ *hisyo* | *ga* **<*gap*>**$_i$ |
| meeting | at | lawyer | ACC | president | TOP | secretary | NOM |

| | | |
|---|---|---|
| *sagasiteiru* | *to*] | *itta*]. |
| look.for | COMP | said |

The sentence in the canonical condition (CC) has a typical Japanese word order (TOP – NOM – **ACC**). In the scrambled condition (SC), the object NP moves from an embedded clause to the matrix clause, crossing two clause boundaries (**ACC**$_i$ $[_{CP}$ TOP $[_{CP}$ NOM <gap>$_i$]])). The results of ERP experiment showed the following three findings. (i) A sustained anterior negativity (sAN) was elicited at the first two NPs in the SC compared to the CC. (ii) Subsequently, a widely distributed positive component (with left fronto-temporal maximum) was observed at the subject *hisyo ga* ‘secretary MON’ in the SC. (iii) Finally, at the second VP *itta* ‘said’, a bilateral negative component was observed in the anterior area in SC compared to the CC. (See Koizumi’s chapter in this volume on scrambling.)

Hagiwara et al. (2007) interpreted the sAN effect as reflecting the storage cost to hold an NP-filler in working memory. In previous studies on filler-gap dependencies of the wh-construction, however, the negativity started at the word following the wh-phrase. That is, no effects were observed at the wh-filler itself. Contrary to this, the sAN in Hagiwara et al. (2007) started at the scrambled-NP filler position. The authors suggest the possibility that the NP-filler may be easier to detect and may be more prone to eliciting a sustained negativity than the wh-filler. However, note that the comparison between the scrambled-NP (*bengosi o*) in SC and the topic word (*syatyoo wa*) in CC is the comparison between two different phrases. Actually, the authors have correctly mentioned that the sAN effect at the word next to the scrambled-NP would be the genuine component of the sAN. After the effect of the scrambled-NP disappeared, we actually start to memorize and store this NP in working memory. Thus, the anterior negativity at the scrambled-NP position would be a phasic effect, which would reflect that we recognize the dislocated object as being scrambled, although the negativity may reflect only the lexical-semantic difference of compared items.

The positivity at the pre-gap NP position *hisyo ga* ‘secretary NOM’ was interpreted as the “integration P600”, which indexes the cost of the structural integration as was observed in the case of the integration of the wh-filler and its gap. The P600

in English wh-constructions (Phillips, Kazanina and Abadac 2005) had been primarily observed at the posterior region, while the P600 in this Japanese scrambled sentence was observed at a left fronto-temporal maximum. The discrepancy between the topographical distributions was claimed to be due to the difference in word order between English and Japanese. The process of integration in English occurs at the pre-gap verb position because the thematic roles are assigned by a verb. On the other hand, integration is performed at the pre-gap NP position in Japanese where thematic-role assigning verbs do not appear yet. However, it is not clear why this difference of position between pre-gap verb and pre-gap NP integrations produced a topographic difference (posterior vs. left fronto-temporal).

At the position of the matrix verb *itta* 'said' a phasic anterior negativity (pAN) was elicited, while at the embedded verb (*sagasiteiru* 'look for') no such ERP effect was observed. The authors interpreted this pAN effect as reflecting thematic-role assignment and subsequent compositional interpretation of sentential meaning. In the stimuli sentences, the checking of grammatical relations in all the three NPs must take place at the matrix verb, not at the embedded verb that can take only two arguments.

Four different ERP components are suggested to reflect four functionally distinct subprocesses in Japanese scrambled constructions. (i) The pAN at the scrambled-NP position may reflect the process of recognizing the scrambled element. (ii) The sAN reflects the process of holding NP-fillers in working memory. (iii) The structural integration process is indexed by the P600 at the pre-gap NP position. (iv) The pAN at the main verb position reflects the thematic-role assignment and compositional interpretation of the whole meaning of the sentence.

### 5.3.4 Summary of integration processes in filler-gap dependency

The findings from wh-constructions (English and German) and scrambled sentences (Japanese) can be summarized as follows.

(24) Filler-gap dependencies indexed by the ERP components in English, German, and Japanese (pAN = phasic AN; s(L)AN = sustained (L)AN; <word> is an adjacent word next to the dependent element).

a. Wh-constructions in English (Phillips, Kazanina and Abadac 2005)

| wh-element | <word> | **Verb <gap>** |
|---|---|---|
| | sAN | P600 |

b. Wh-constructions in German (Fiebach, Schlesewsky and Friederici 2001)

| wh-element | <word> | **NP <gap>** | Verb |
|---|---|---|---|
| | sLAN | P600 | no P600 |

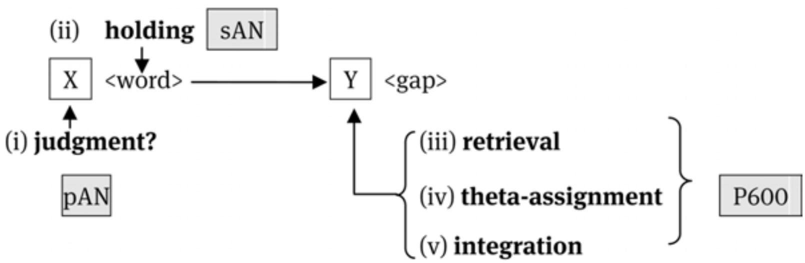
c. Scrambled sentences in Japanese (Hagiwara et al., 2007)

scrambled-NP <word> **NP <gap>** Verb

pAN sAN P600 pAN

Here, let us assume there is a long-distance dependency relationship between a dependent (unassociated) element "X" and a depended (associating) element "Y", then we can propose a rough design of architecture for processing this dependency as follows.

(25) The process of long-distance dependency




On encountering an element X, we have to judge whether X is a dependent element that requires a depended element Y in the downstream of the sentence. This judgment process may be indexed by a phasic anterior negativity (pAN) in the Japanese example. However, since these pAN effects were obtained by the comparison of different lexical items, we need some caution. Once we have judged X is a dependent element, X has to be held in working memory. This holding cost is indexed by a sustained anterior negativity (sAN) without exception. At the point where the depended element Y appears, three functionally distinct subprocesses are proposed, although they are all indexed by the P600. Since the dependent element X has been held in working memory, X has to be retrieved from the memory for the purpose of integration. Thus the retrieval is a pre-integration process. In a wh-construction of head-initial languages such as English, thematic-role assignment occurs at the pre-gap verb position. In head-final constructions, however, the integration occurs at the pre-gap NP position before the thematic-role assigning verb appears. Thus, this thematic-role assignment process for integration depends on the nature of the construction in question. Finally, the overall integration process between X and Y is performed to reach the final stage of sentence comprehension.

Considering what kind of constructions in which language elicit what kind of ERP effects, we proposed the sketches in (25). Although they are very rough, they

can give us an outline to examine the processing of dependency relation between two discrete elements located in distance.[6]

## 5.4 Summary of Section 5

In order to illustrate the underlying mechanism of language processing, in Section 5 we examined three types of linguistic information sources that can impact syntactic aspects of language processing with the modulation of the P600 effects. In Section 5.1, we observed that ungrammatical wh-questions in Japanese elicited the "repair P600" that is interpreted as reflecting the syntactic repair process to try to construct a coherent message from ungrammatical constructions. Furthermore, case-assignment violations in multiple languages exhibit the biphasic pattern of the LAN (or N400) and the P600, indicating that the morpho-syntactic violation is processed in a two-step fashion: morphological violation checking and syntactic repairing. In Section 5.2 we examined garden-path sentences and found that non-preferred constructions elicit the "revision P600" effect compared to preferred counterparts even though both are grammatical. It was demonstrated that the garden-path (reanalysis) effect was caused by the preference information embodied in a verb of a head-initial language like English and in a structural simplicity measure of a head-final language like Japanese. Section 5.3 was devoted to discussions on the "integration P600" effects of long-distance dependency across various languages and constructions. There appeared two basic ERP components. One is a sustained anterior negativity related to working memory to hold a dependent element. The other is the P600 indexing the cost of integration of two distinct elements in a sentence. Some sub-processes with attendant ERP components were also proposed to explain the more detailed mechanism of integration.

**6** In this subsection we observed long-distance filler-gap relationships where the filler is overt (visible) and the gap is covert (invisible). Investigating Japanese wh-questions that do not involve a gap, Ueno and Kluender (2009) failed to find integration P600 effects. Examining an association process of numeral classifiers and corresponding NPs, however, Yasunaga and Sakamoto (2007) observed integration P600 effects. We need to examine long-distance dependencies involving various types of constructions: for example, negative polarity items (cf. Bise and Sakamoto 2010) such as *Taroo* ***sika*** *hon o yoma-****nai***. 'Only Taroo reads a book.', and concessive clauses (cf. Tateyama et al. 2012) such as ***Tatoe*** *ame ga fut-****temo*** *dekakeru*. 'Even if it rains, I will go out.' Accumulation of experimental research would reveal whether the integration process between various elements in Japanese is governed by a more general principle.

So far, we examined the integration processes where a "dependent" (unassociated) element (e.g., a scrambled NP) precedes a "depended" (associating) element (e.g., a gap). However, there are some interesting constructions where the depended element precedes the dependent element. In this case the integration process proceeds in a backward fashion. Investigating Japanese relative clauses, Ueno and Garnsey (2008) reported that object relative clauses elicited larger P600 effects than subject relatives. Examining floating classifier sentences, Yasunaga, Oishi and Sakamoto (2007) also observed the integration P600 effect.

Considering the discussions in this section, we propose the second tentative answer to the three questions (WHAT, WHEN, and HOW) raised in the Introduction as follows. We use syntactic information sources such as ungrammaticality, garden-path, and long-distance dependency during the time window of around 600 ms in an expectancy-driven way. Depending on the nature of the construction in question, the P600 can be preceded by a negativity such as the phasic (L)AN, the sustained (L) AN, or the N400. Since we have observed processing of semantic and syntactic information in Sections 4 and 5, respectively, the next section reconsiders the relationship between semantic and syntactic processing.

# 6 Syntax-semantics interactions and ERPs

We observed that distinct ERP components reflect different types of linguistic processing. In other words, the brain honors the dissociation between syntax and semantics. In order to accomplish full understanding of a sentence, however, these two sources of linguistic information must ultimately be unified. This section first presents some empirical evidence of dynamic interaction between syntactic and semantic processing. Then, we examine two models to account for this dynamic interaction.

## 6.1 Challenges to syntax-semantics dissociations: the semantic P600

As we have seen in previous sections, the one-to-one mapping between linguistic domains (semantics and syntax) and ERP components (N400 and P600, respectively) appears to be relatively robust in both English and Japanese. Recent studies, however, have begun to show that certain kinds of semantic/thematic anomalies failed to elicit (predicted) N400 effects but rather elicited P600 effects. For example, consider the following examples form Kuperberg et al. (2003).

(26) a. Non-anomalous control sentences
*For breakfast the **boys** would only **eat** toast and jam.*

b. Thematic-role animacy violation sentences
*For breakfast the **eggs** would only **eat** toast and jam.*

In (26b) syntactic cues unambiguously indicate that the verb *eat* assigns the Agent role to the inanimate noun *egg*. The situation is similar to the selectional restriction violation that elicits a typical N400 effect. Remember the example (1b): “*The cat will bake the food…*” that we discussed in Section 3. Based on our knowledge of lexical-semantic properties, we know that cats cannot bake the food and eggs

cannot eat. Thus we predict that (26b) would elicit an N400 effect. Contrary to this prediction, Kuperberg et al. (2003) observed a robust P600 effect but no N400 enhancement in response to the verb *eat* in the animacy violation condition (26b) compared to the normal counterpart (26a). Similar results to this have been obtained not only in English but also in several other languages, with many related, but different types of manipulations (for review, see Bornkessel-Schlesewsky and Schlesewsky 2008; Kuperberg 2007). These kinds of P600 effects have been dubbed "semantic P600s", due to the fact that these effects had been assumed to be caused by semantic incongruity.

At first glance, such findings are problematic because semantic anomalies elicited P600 effects that have long been assumed to be indices of syntactic processes such as syntactic repair, syntactic revision and syntactic integration as we have observed in Section 5. What is specific to the animacy violation in (26b), however, is that the semantic/thematic information suggests that this inanimate noun is a highly plausible Theme of the verb. Since eggs cannot "eat" but can "be eaten", a reversed assignment of Agent and Theme roles would yield a semantically and pragmatically plausible interpretation. Thus, Kuperberg et al. (2003), suggest the possibility that the P600 was elicited by "an online attempt to structurally repair and make sense of the sentences by reassigning the thematic role of the NP that preceded the critical verb from 'agent' to 'theme'" (p. 117). That is, the P600 effect in this reversal thematic-role situation is interpreted to reflect the syntactic repair process as have been observed in various syntactic constructions discussed in previous sections especially in Section 5.1. In this "syntactic repair" hypothesis on semantic P600 effects, the functional dissociation between syntax and semantics is maintained as such because these effects do not reflect the putative semantic processes but the syntactic processes (see also Hoeks, Stowe and Doedens 2004; Kim and Osterhout 2005).

However, van Herten et al. (2005) propose a "conflict" hypothesis, which distinguishes the "plausibility heuristic processes" (i.e., the interpretation that fits world knowledge) and the "algorithmic syntactic analysis". We deal with two different possible thematic-role interpretations, one induced by the plausibility heuristics and the other constructed by the syntactic algorithm. The semantic P600 is considered to reflect this conflict between these two thematic-role interpretations in thematic-role reversal conditions (See also Bornkessel-Schlesewsky and Schlesewsky 2008; Kuperberg 2007).

The findings of the semantic P600 were obtained mostly from the investigation of Indo-European languages which have typological, historical, and constructive similarities. There is very little research on the semantic P600 effects from non-Indo-European languages (cf. Ye and Zhou 2008). In order to assure that the semantic P600 phenomena is an important clue to clarify the universal property of processing architecture, we need to accumulate more findings across various languages. Here, we examine a semantic P600 effect elicited in response to semantically anomalous but

thematically reversal sentences in Japanese. Consider the following examples from Oishi and Sakamoto (2009).

(27) a. Grammatical and semantically non-anomalous sentences

| | | | | |
|---|---|---|---|---|
| *Umaretate* | *no* | ***imomusi*** | *ga* | *midorino* |
| new.born | GEN | green.caterpillar | NOM | green |

| | | |
|---|---|---|
| ***happa*** | *ni* | ***kaziritui-ta.*** |
| leaf | DAT | bit |

'A new-born green caterpillar bit a green leaf.'

b. Grammatical but semantically anomalous sentences

| | | | | |
|---|---|---|---|---|
| ?*Umaretate* | *no* | ***imomusi*** | *ga* | *midorino* |
| new.born | GEN | green.caterpillar | NOM | green |

| | | |
|---|---|---|
| ***happa*** | *ni* | ***kazirituk-are-ta*.** |
| leaf | DAT | was bitten |

'A new-born green caterpillar was bitten by a green leaf.'

Although both (27a) and (27b) are syntactically well-formed, (27b) is semantically anomalous: a green caterpillar cannot be bitten by a leaf. If we assume the traditional dissociation between syntax and semantics, this semantic incongruity should have elicited N400 effects. Contrary to this expectation, a P600 effect was observed as shown in Figure 11, replicating the previous studies described above.

Following the above discussion of the two possible hypotheses concerning the semantic P600 effects, there could be two possible interpretations on this result

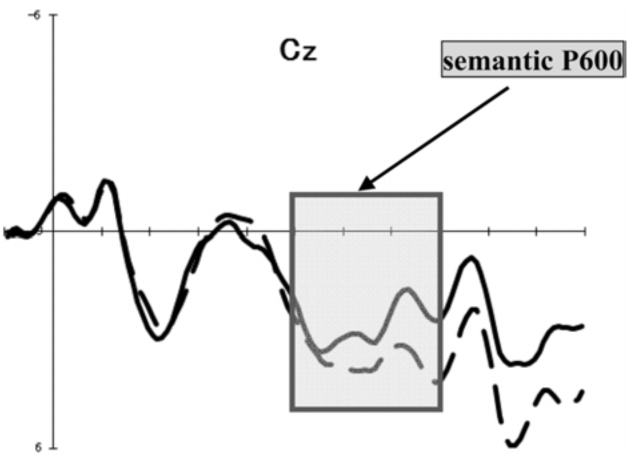



**Figure 11:** Grand Average ERPs at Cz elicited by the matrix verbs. ERPs to non-anomalous sentences (27a) are plotted with solid lines, those to semantically anomalous sentences (27b) are plotted with large dotted lines. Negativity is plotted up.

from Japanese data. The syntactic "repair" hypothesis assumes that the semantic association between *imomusi* 'green caterpillar', *happa* 'leaf', and *kazirituku* 'bite' is so strong that the semantic information overrides syntactic information and we misdiagnose this sentence as "syntactically" anomalous. We may incorrectly assume that the cause of the anomaly is the presence of the passive form *kazirituk-are-ta* 'was bitten' instead of what is considered to be the correct active form *kaziritui-ta* 'bit'. Second, the "conflict" hypothesis assumes that the semantic association between content words in the sentence is strong enough for the semantic/pragmatic processing to output the pragmatically plausible interpretation: "*The green caterpillar bit the leaf*". This interpretation fits well into our world knowledge. On the other hand, algorithmic syntactic parsing tells us that the sentence "*The green caterpillar was bitten by the leaf*" describes a pragmatically anomalous situation, i.e., "*The leaf bit the green caterpillar*". As such, the semantic P600 is interpreted to reflect the processing cost emerging from the conflict between semantic/pragmatic and syntactic processing streams. Of course, more data are necessary to evaluate the validity of these two possible interpretations.

These types of effect in which the P600 is elicited by semantic/thematic anomalies have attracted considerable attention in the field of psycholinguistics as they challenge some of the most foundational and well-accepted parsing models within the field. In fact, the idea that semantic outputs challenge syntactic outputs in the absence of syntactic ambiguity is crucially problematic for "syntax-first" models (Ferreira and Clifton 1986; Frazier and Rayner 1982; Friederici 2002). These models assume that syntactic cues alone are in control of first-pass analysis and that semantic information does not exert an influence until later in the parsing process. The "semantic challenge to syntax" idea is also problematic for a weak version of "interactive" or constraint-based models (MacDonald, Pearlmutter and Seidenberg 1994; Trueswell, Tannenhaus and Garnsey 1994). Interactive models allow semantic information to have an impact on first-pass analysis, but only when ambiguity arises due to a lack of decisive syntactic information. These weak interactive models assume that syntactic cues, when unambiguous, will dominate first-pass structure-building operations. A strong version of an interactive model such as the "immediacy" model (Hagoort 2007), however does permit us to make use of any source of information as soon as it becomes available and as such is consistent with the conflict hypothesis of semantic P600 effects.

The discovery of semantic P600 effects provides electro-physiological evidence that syntactic processes do not always take priority over semantic processes during language comprehension. In fact, these two processes appear to be interacting dynamically in our brain. Without the fine-grained temporal resolution of ERPs, these dynamic processes would not have been observable. In the next subsection, we will briefly review the syntax-first and interactive (immediacy) models.

## 6.2 Models for syntax-semantics interactions

Concerning the interaction of various types of linguistic processes, as mentioned in the above section, there are two main proposals: "syntax-first" models and "interactive" models. The former assumes that autonomous syntactic processes precede semantic processes in such a way that we develop the simplest syntactic structure independent of semantic information (or, indeed, any other source of information), while the latter claim that lexical-semantic and more global semantic information influence syntactic structure building operations (as well as other factors, such as frequency, etc.).

Friederici (2002: 79) proposed a syntax-first three-phase model of sentence comprehension, which has characteristics of modular, serial, and hierarchical architecture. Figure 12 is a modified and simplified representation of the proposed model.

Phase 0 (0–100 ms) is not relevant to linguistic processing but rather to general perceptual processes such as acoustic analysis of incoming speech input. During Phase 1 (100–300 ms) structure building operations occur based on word-category information alone. Morpho-syntactic and lexical-semantic processes occur in Phase 2 (300–500 ms) resulting in thematic-role assignment. During Phase 3 (500–1000 ms), integration of different types of information takes place, and the comprehension system arrives at a final interpretation of the sentence. The process of building the syntactic structure is autonomous and precedes semantic processes in Phases 1 and 2 respectively, with these processes only interacting in Phase 3 in order to output a final representation.

Friederici's (2002) syntax-first model claims that the earliest stage of processing is purely syntactic in nature so that an initial syntactic structure is formed on the basis of word-category information alone. Only in later stages of processing do different outputs of each process interact in this hierarchical model. Contrary to this syntax-first model, a strong version of interactive models argues that we utilize any information as soon as it becomes available. For example, the MUC (Memory, Unification, and Control) model proposed by Hagoort (2007) claim that "the different processing levels (phonological, syntactic, semantic/pragmatic) operate in parallel, and, to some degree independent. Where necessary, cross-talk takes place, which is again characterized by the immediacy principle. That is, cross-talk takes place on a more or less moment-to-moment basis (p. 285)".

The modular, serial, and hierarchical characteristics of the syntax-first model helps us to consider the time course of language processing, although this model suffers from explaining free interactions of various types of language processes. The strong interaction model can explain the free "cross-talk" between various processes, although it is difficult to specify when and how the cross-talk occurs. As is often the case, these two models of parsing are both supported by the results of ERP experiments. Undoubtedly we need much more accumulation of experimental

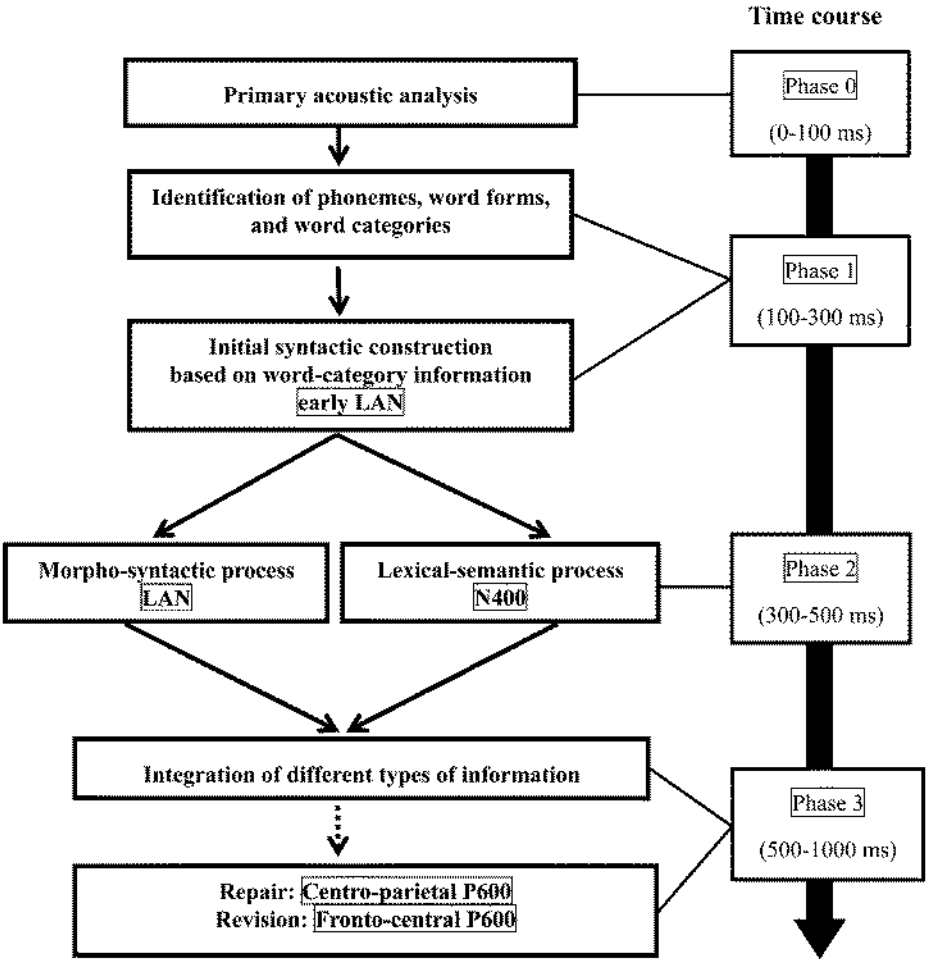



**Figure 12:** The time course of language processing and the relevant ERP components based on Friederici (2002) with some revision and simplification.

and theoretical studies to clarify the architecture of the comprehension system, including studies that attempt to identify the locus of the interaction of different information sources. In this sense, the ERP studies on Japanese, as discussed in previous sections, can contribute to this enterprise by providing data on a language with very different properties from the Indo-European languages mostly examined thus far in the literature.

# 7 Concluding remarks and future direction

We have reviewed various Japanese ERP studies with some comparison of similar studies in English and German. We found very interesting similarities and contrasts among these languages. For instance, in the discussion of truth-value verification constructions the head-final nature of Japanese helped to eliminate a possible confounding of two distinct processing factors (Section 4.2.1). In a head-initial language like English, verb information is assumed to determine a preferred construction in garden-path phenomena, while in a head-final language like Japanese verb information comes last so that syntactic simplicity metrics are considered to determine the preference before the sentence final verb appears (Section 5.2). In a long-distant dependency, the exact point of integration is crucially different between the head-initial (verb position) and head-final (NP position) constructions (Section 5.3.1). These findings deserve special mention in the sense that they reveal the importance of cross-linguistic research in order to clarify the universality and specificity of language processing.

The overall discussion in this chapter can be summarized in the following illustration. We posited three questions (WHAT, WHEN, and HOW) in the Introduction. In order to answer the question HOW, we observed two types of semantic process (i: Plausibility and ii: Relatedness) and three types of syntactic process (i: Repair, ii: Revision, and iii: Integration). As a possible illustration for the question WHAT, we examined two types of semantic information sources (1: Cloze probability and 2:

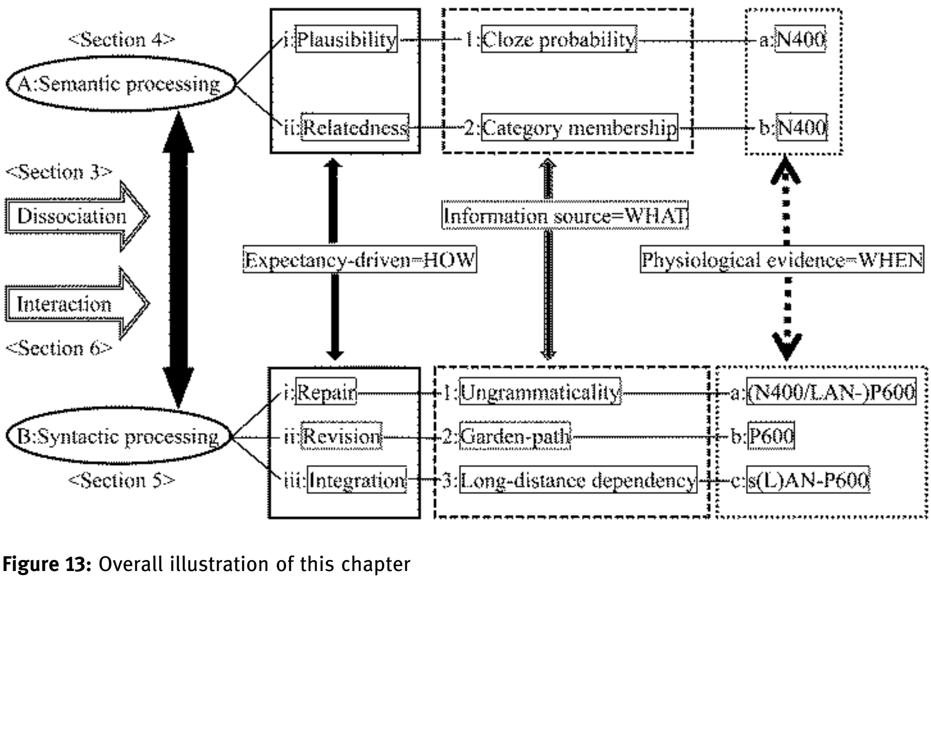



**Figure 13:** Overall illustration of this chapter

Category membership) and three types of syntactic information sources (1: Ungrammaticality, 2: Garden-path, and 3: Long-distance dependency). For the purpose of answering the question WHEN, we explored what kind of ERP component reflects the underlying language processes. Note that the "processes" assumed here are theoretical constructs, which cannot be proved to exist without some kind of "evidence". Discussions in this chapter were developed to the functional relationship between the linguistic (syntactic and semantic) processes and the corresponding physiological evidence (i.e., ERP index such as the N400 and the P600).

Now, due is the final answer to the three questions: WHAT, WHEN, and HOW. In the time course of language comprehension, we use semantic information sources during the time window of around 400 ms and syntactic information sources around 600 ms. The semantic information sources include semantic plausibility and semantic relatedness, while the syntactic sources involve ungrammaticality, garden-path, and long-distance dependency. We handle these information sources in a expectancy-driven way. The physiological evidence reviewed here reveals that expectations are generated at multiple linguistic levels, and that these different varieties of expectation can be mapped onto distinct ERP components. These processes can be independent or interactive, depending on the specific challenges presented by the input being processed. It is just the type of results reviewed here that will be critical in constructing a comprehensive theory of cross-linguistic sentence processing as the field moves forward. We believe that it is important to continue to pursue an understanding of how the brain processes the Japanese language, as these types of studies can reveal much about cross-linguistic similarities and differences in terms of how the brain processes language. Studying Japanese with ERPs is indeed an important endeavor, and one that can make contributions to the understanding of human language in general.

## Acknowledgments

I gratefully thank Chris Barkley and Mieko Ueno for giving me precious comments to improve previous versions of this article. I also thank Jun'ichi Katayama, Yoan Kim, Hiroshi Nittono, Hiroaki Oishi, Nicola Strambini, Shugo Suwazono, Yuki Tateyama, Masataka Yano, and Daichi Yasunaga for their helpful comments. I would like to thank the editor Mineharu Nakayama for giving me the opportunity to contribute to this volume. Part of this work is supported by JSPS KAKENHI Grant Number 25244018 and 22222001.

# References


Anaki, David, Miriam Faust and Shlomo Kravetz. 1998. Cerebral hemispheric asymmetries in processing lexical metaphors. *Neuropsychologia* 36(4). 353–362.

Arao, Hiroshi, Shugo Suwazono, Tsutomu Sakamoto and Tsutomu Nakada. 2007. ERP correlates of the processing of object-verb integration in Japanese. In Tsutomu Sakamoto (ed.), *Communicating skills of intention*, 319–336. Tokyo: Hituzi Syobo.

Bise, Yu and Tsutomu Sakamoto. 2010. Examination of Typing Mismatch Effect in processing of Japanese sika-nai construction. *IEICE Technical Report* 110(163). 31–36.

Bise, Yu and Tsutomu Sakamoto. 2011. *Nihongo bunshori niokeru nijūtaikaku seiyaku no shinriteki jitsuzaisei nitsuite* [Remarks on the psycholinguistic reality of the double-o constraint in Japanese sentence processing]. *SIG-SLUD-B101-07*. 29–34.

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky. 2008. An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews* 59(1). 55–73.

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky. 2009. *Processing syntax and morphology: A neurocognitive perspective*. New York: Oxford University Press.

Brown, Colin M. and Peter Hagoort. 1993. The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience* 5(1). 34–44.

Brownell, Hiram H., Heather H Potter, Diane Michelow and Howard Gardner. 1984. Sensitivity to lexical denotation and connotation in brain-damaged patients: A double dissociation? *Brain and Language*. 22(2). 253–265.

Coulson, Seana, Jonathan W. King and Marta Kutas. 1998. Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes* 13(1). 21–58.

Chwilla, Dorothee J., Colin M. Brown and Peter Hagoort. 1995. The N400 as a function of the level of processing. *Psychophysiology* 32. 274–285.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Díaz, Begoña, Núria Sebastián-Galles, Kepa Erdocia, Jutta L. Mueller and Itziar Laka. 2011. On the cross-linguistic validity of electrophysiological correlates of morphosyntactic processing: A study of case and agreement violation in Basque. *Journal of Neurolinguistics* 24. 357–373.

Ferreira, Fernanda and Charles Clifton Jr. 1986. The independence of syntactic processing. *Journal of Memory and Language* 25. 348–368.

Fiebach, Christian J., Matthias Schlesewsky and Angela D. Friederici. 2001. Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *Journal of Psycholinguistic Research* 30. 321–338.

Fischler, Ira, Paul A. Bloom, Donald G. Childers and Salim E. Roucos. 1983. Brain potentials related to stages of sentence verification. *Psychophysiology* 20(4). 400–409.

Frazier, Lyn and Charles Clifton, Jr. 1998. Sentence reanalysis, and visibility. In Janet D. Fodor and Fernanda Ferreira (eds.), *Reanalysis in sentence processing*, 143–176. Dortrecht: Kluwer Academic Publisher.

Frazier, Lyn and Janet D. Fodor 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6. 291–325.

Frazier, Lyn and Keith Rayner. 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally-ambiguous sentences. *Cognitive Psychology* 14. 178–210.

Friederici, Angela D. 2002. Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences* 6(2). 78–84.

Friederici, Angela D. and Stefan Frisch 2000. Verb argument structure processing: The role of verb-specific and argument-specific information. *Journal of Memory and Language* 43. 476–507.

Friederici, Angela D., Karsten Steinhauer and Stefan Frisch. 1999. Lexical integration: Sequential effects of syntactic and semantic information. *Memory and Cognition* 27. 438–453.
Frisch, Stefan and Matthias Schlesewsky. 2001. The N400 reflects problems of thematic hierarchizing. *Neuroreport: For Rapid Communication of Neuroscience Research* 12. 3391–3394.
Frisch, Stefan and Matthias Schlesewsky. 2005. The resolution of case conflicts from a neurophysiological perspective. *Cognitive Brain Research* 25. 484–498.
Giora, Rachel. 1999. On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics* 31. 919–929.
Hagiwara, Hiroko. 2015. Language acquisition and brain development: cortical processing of a foreign language. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.
Hagiwara, Hiroko, Heizo Nakajima, Kazuyuki Nakagome, Satoru Takazawa, Osamu Kanno, Kenji Itoh and Ichiro Koshida. 2000. ERP manifestations of processing syntactic dependencies in hierarchical structures of language: Time course and scalp distribution. In Kazuko Inoue (ed.) *Researching and verifying an advanced theory of human language*, 519–545. Kanda University of International Studies
Hagiwara, Hiroko, Takahiro Soshi, Masami Ishihara and Kuniyasu Imanaka. 2007. A topographical study on the event-related potential correlates of scrambled word order in Japanese complex sentences. *Journal of Cognitive Neuroscience* 19. 175–193.
Hagoort, Peter. 2007. The memory, unification, and control (MUC) model of language. In Tsutomu Sakamoto (ed.), *Communicating skills of intention*, 259–291. Tokyo: Hituzi Syobo.
Hagoort, Peter, Colin M. Brown and Lee Osterhout. 1999. The neurocognition of syntactic processing. In Colin M. Brown and Peter Hagoort (eds.) *The neurocognition of language*, 273–316. Oxford UK: Oxford University Press.
Hoeks, John C.J., Laurie A. Stowe and Gina Doedens. 2004. Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research* 19. 59–73.
Holcomb, Phillip J. 1988. Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language* 35. 66–85.
Joo'o, Hakutaroo. 1996. Perception of pitch, from point of view of experimental linguistics. *Studies in Language and Literature* 30. 15–35. University of Tsukuba.
Kaan, Edith, Anthony Harris, Edward Gibson and Phillip Holcomb. 2000. The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15 (2). 159–201.
Kaan, Edith and Tamara Y. Swaab. 2003. Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience* 15(1). 98–110.
Katayama, Jun'ichi. 1995. *Imiteki na kitai no shinriseirigaku* [Psychophysiology of semantic expectancy]. Tokyo: Taga Syuppan.
Katayama, Jun'ichi, Yo Miyata and Akihiro Yagi. 1987. Sentence verification and event-related brain potentials. *Biological Psychology* 25. 173–185.
Kim, Albert and Osterhout, Lee. 2005. The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language* 52(2). 205–225.
King, Jonathan W. and Marta Kutas. 1995. Who did what and when?: Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience* 7. 376–95.
Kluender, Robert and Marta Kutas. 1993a. Bridging the gap: Evidence from ERPs on the processing of unbound dependencies. *Journal of Cognitive Neuroscience* 5(2). 196–214.
Kluender, Robert and Marta Kutas. 1993b. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8(4). 573–633.
Kobayashi, Yuki, Ichiro Kanamaru, Yoko Sugioka and Takane Ito. 2007. An ERP study of the processing of case marker violations in Japanese. *IEICE Technical Report* 107(138). 45–50.

Koizumi, Masatoshi. 2015. Experimental syntax: word order in sentence processing. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Koso, Ayumi, Hiroko Hagiwara and Takahiro Soshi. 2007. Event-related brain potentials associated with scrambled Japanese ditransitive sentences. In Tsutomu Sakamoto (ed.), *Communicating skills of intention*, 337–352. Tokyo: Hituzi Syobo Publishing.

Koso, Ayumi, Shiro Ojima and Hiroko Hagiwara. 2011. An event-related potential investigation of lexical pitch-accent in auditory Japanese. *Brain Research* 1385. 217–228.

Kuperberg, Gina R. 2007. Neural mechanisms of language comprehension: challenges to syntax. *Brain Research* 1146. 23–49.

Kuperberg, Gina R., Tatiana Sitnikova, David Caplan and Phillip J. Holcomb. 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research* 17(1). 117–129.

Kusumi, Takashi. 1988. Comprehension of synesthetic expressions: Cross-modal modifications of sense adjectives. *The Japanese Journal of Psychology* 58(6). 373–380.

Kutas, Marta and Kara D. Federmeier. 2011. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology* 62. 621–647.

Kutas, Marta and Steven A. Hillyard. 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207. 203–205.

Kutas, Marta and Steven A. Hillyard. 1989. An electrophysiological probe of incidental semantic association. *Journal of Cognitive Neuroscience* 1(1). 38–49.

Kutas, Marta, Timothy E. Lindamood and Steven A. Hillyard. 1984. Word expectancy and event-related brain potentials during sentence processing. In Sylvan Kornblum and Jean Requin (eds.). *Preparatory states and processes*, 217–237. Hillsdale, New Jersey: Lawrence Erlbaum.

MacDonald, Maryellen C., Neal J. Pearlmutter and Mark S. Seidenberg. 1994. The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101(4). 676–703.

Miyagawa, Shigeru. 1989. *Syntax and Semantics 22: Structure and case marking in Japanese*. New York: Academic Press.

Molinaro, Nicola, Horacio A. Barberb and Manuel Carreiras. 2011. Grammatical agreement processing in reading: ERP findings and future directions. *Cortex* 47. 908–930.

Mueller, Jutta L., Masako Hirotani and Angela D. Friederici. 2007. ERP evidence for different strategies in the processing of case markers in native speakers and non-native learners. *BMC Neuroscience* 8(18). 1–16.

Nakagome, Kazuyuki, Satoru Takazawa, Osamu Kanno, Hiroko Hagiwara, Heizo Nakajima, Kenji Itoh and Ichiro Koshida. 2001. A topographical study of ERP correlates of semantic and syntactic violations in the Japanese language using the multi-channel EEG system. *Psychophysiology* 38. 304–315.

Nakao, Mizuki and Makoto Miyatani. 2007. Dissociation of semantic and expectancy effects on N400 using Neely's version of semantic priming paradigm: N400 reflects post-lexical integration. In Tsutomu Sakamoto (ed.), *Communicating skills of intention*, 201–212. Tokyo: Hituzi Syobo.

Nashiwa, Hitomi, Mizuki Nakao and Makoto Miyatani. 2007. Interaction between semantic and syntactic processing in Japanese sentence comprehension. In Tsutomu Sakamoto (ed.), *Communicating skills of intention*, 311–318. Tokyo: Hituzi Syobo.

Neely, James H. 1977. Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General* 106. 226–254.

Neely, James H. 1991. Semantic priming effects in visual word recognition: A selective review of current findings and theories. In Derek Besner and Glyn W. Humphreys (eds.), *Basic processes in reading: Visual word recognition*, 264–336. Hillsdale, NJ: Lawrence Erlbaum.

Oishi, Hiroaki and Tsutomu Sakamoto. 2009. Immediate interaction between syntactic and semantic outputs: evidence from event-related potentials in Japanese sentence processing. Poster presented at The annual CUNY Sentence Processing Conference. 2009. UC Davis, 27 March.

Oishi, Hiroaki, Daichi Yasunaga and Tsutomu Sakamoto. 2007. Revision process in Japanese sentence processing: Evidence from event-related brain potentials. In Tsutomu Sakamoto (ed.), *Communicating skills of intention*, 367–381. Tokyo: Hituzi Syobo.

Osterhout, Lee and Phillip J. Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31. 785–806.

Osterhout, Lee and Phillip J. Holcomb. 1993. Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes* 8(4). 413–437.

Osterhout, Lee, Phillip J. Holcomb and David A. Swinney. 1994. Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology* 20 (4). 786–803.

Osterhout, Lee and Kayo Inoue. 2007. What the brain's electrical activity can tell us about language processing and language learning. In Tsutomu Sakamoto (ed.), *Communicating skills of intention*, 293–309. Tokyo: Hituzi Syobo.

Osterhout, Lee and Janet Nicol. 1999. On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes* 14. 283–317.

Phillips, Colin, Nina Kazanina and Shani H. Abada. 2005. ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22. 407–428.

Sadakane, Kumi and Masatoshi Koizumi. 1995. On the nature of the "dative" particle *ni* in Japanese. *Linguistics* 33. 5–33.

Sakai, Yumi, Kazuki Iwata, Jorge Rira, Xiaohong Wan, Satoru Yokoyama, Yoshiteru Shimoda, Ryuta Kawashima, Kei Yoshimoto and Masatoshi Koizumi. 2006. *Jishō kanren den'i-de miru meishi to josūshino shōgō purosesu: Imiteki shorika bun'pōteki shorika*. [An ERP study of the integration process between a noun and a numeral classifier: Semantic or morphosyntactic?] *Ninchi Kagaku* 13(3). 443–454.

Sakamoto, Tsutomu. 1983. Towards systematic treatment of synaesthetic metaphor. *Proceedings of The Kansai Linguistic Society* 3. 95–104.

Sakamoto, Tsutomu , Kana Matsuishi, Hiroshi Arao and Junri Oda. 2003. An ERP study of sensory mismatch expressions in Japanese. *Brain and Language* 86(3). 384–394.

Sturt, Patrick and Matthew W. Crocker. 1996. Monotonic syntactic processing: A cross-linguistic study of attachment and reanalysis. *Language and Cognitive Processes* 11 (5). 449–494.

Takazawa, Satoru, Nobuaki Takahashi, Kazuyuki Nakagome, Osamu Kanno, Hiroko Hagiwara, Heizo Nakajima, Kenji Itoh and Ichiro Koshida. 2002. Early components of event-related potentials related to semantic and syntactic processes in the Japanese language. *Brain Topography* 14. 169–77.

Tamaoka Katsuo, Nobuhiro Saito, Sachiko Kiyama, Kalinka Timmer and Rinus G. Verdonschot. 2014. Is pitch accent necessary for comprehension by native Japanese speakers? – An ERP investigation. *Journal of Neurolinguistics* 27. 31–40.

Tateyama, Yuki, Yu Bise, Masataka Yano, and Tsutomu Sakamoto. 2012. *"Tatoe-temo" bun no shori nitsuite: Jishōkanrenden'i o shihyōtoshite*. [On the processing of concessive tatoe V-temo clauses in Japanese: An ERP study.] *IEICE Technical Report* 112(145). 25–30.

Trueswell, John C., Michael K. Tanenhaus and Susan M. Garnsey. 1994. Semantic influence on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33. 285–318.

Ueno, Mieko and Susan M. Garnsey. 2008. An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes* 23(5). 646–688.

Ueno, Mieko and Robert Kluender. 2003. Event-related brain indices of Japanese scrambling. *Brain and Language* 86. 243–271.

Ueno, Mieko and Robert Kluender. 2009. On the processing of Japanese wh-questions: An ERP study. *Brain Research* 1290. 63–90.

Ullmann, Stephen. 1951. *The principles of semantics*. Oxford: Blackwell.

van Herten, Marieke, Herman H.J. Kolk, and Dorothee J. Chwilla. 2005. An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research* 22(2). 241–255.

Williams, Joseph M. 1976. Synaesthetic adjectives: A possible law of semantic change. *Language* 52. 461–478.

Winner, Ellen, and Howard Gardner. 1977. The comprehension of metaphor in brain-damaged patients. *Brain*. 100. 717–729.

Yasunaga, Daichi and Tsutomu Sakamoto. 2007. On-line processing of floating quantifier constructions in Japanese: Using event-related brain potentials. *Journal of Japanese Linguistics* 23. 21–34.

Yasunaga, Daichi, Hiroaki Oishi and Tsutomu Sakamoto. 2007. Backward-integration in Japanese: Evidence from event-related brain potentials. In Tsutomu Sakamoto (ed.), *Communicating skills of intention*, 353–365. Tokyo: Hituzi Syobo.

Ye, Zheng and Xiaolin Zhou. 2008. Involvement of cognitive control in sentence comprehension: Evidence from ERPs. *Brain Research* 1203. 103–115.

Koichi Sawasaki and Akiko Kashiwagi-Wood

# 16 Issues in L2 Japanese sentence processing: Similarities/differences with L1 and individual differences in working memory

## 1 Introduction

Sentence processing in a second language (L2) is a relatively new field in psycholinguistics, and its research design and methodologies are similar to approaches to sentence processing in the first language (L1). Researchers of L2 processing try to find out whether L2 processing is similar to or different from L1 sentence processing. If the two types of processing are not identical, then we also investigate whether there are any shared principles or strategies that can explain the mechanism of L2 processing in general, and moreover, why L2 processing needs to be different from L1 processing. These questions lead us to a more fundamental question in psycholinguistics about how an understanding of L2 processing can contribute to theories about cognitive architecture and the brain.

It is difficult to give concrete answers to the questions stated above. First, research on L2 processing is new, and research on L2 Japanese processing is even newer. Only in the past 10 years have researchers started to examine online data for L2 Japanese. Furthermore, Japanese learners' L1s are heavily skewed toward East Asian languages. For example, 62.3% of those who are studying Japanese at universities throughout the world speak Chinese as their L1 and less than 10% of them speak English as their L1 (Japan Foundation 2011: 111). Consequently, the majority of the L2 Japanese findings, especially at the advanced level, are from Chinese L1 or Korean L1 speakers, which makes cross-linguistic comparison difficult (See Tamaoka, this volume, for detailed discussion of L2 Japanese processing by L1 Chinese-speakers).

This chapter discusses the similarities and differences between L1 and L2 sentence processing, and individual differences stemming from working memory, by reviewing the related literature. Previous studies on L2 sentence processing have generally revealed mixed results containing both similarities and differences in their findings, but the emphasis is often placed on the similarities rather than the differences between L1 and L2 processing. Thus, comparing previous findings in terms of the similarities as well as the differences is worthwhile because it will help identify problems for future studies to address. Moreover, working memory is believed to influence sentence processing and explain individual differences in processing and comprehension performance; however, only a few studies have been conducted on working memory and L2 Japanese. Thus, reviewing recent research on this topic in L2 Japanese broadens our understanding of this field.

This chapter is organized as follows. In Section 2, we discuss the similarities and differences between L1 and L2 Japanese processing by focusing on relative clause processing and incremental processing. Section 3 discusses how individual differences in working memory influence L2 Japanese processing. Finally, Section 4 provides concluding remarks and ideas for future research.

# 2 Similarities and differences between L1 Japanese and L2 Japanese processing

The question of whether L2 processing is the same as L1 processing is a fundamental issue in the L2 processing literature. Previous research findings show mixed results. Many studies indicate that although L2 Japanese beginning learners start out with a nonnative-like processing manner, they become native-like as their proficiency improves. On the other hand, some other research indicates that L2 Japanese learners cannot always perform native-like processing even after their proficiency reaches the advanced level. In this section, we will discuss the similarities and differences between L1 and L2 Japanese processing by first looking at the processing of relative clauses and then addressing incremental processing.

## 2.1 Similarities and differences in the processing of relative clauses

Relative clauses are one of the most widely studied topics in both L1 and L2 processing. Many L1 studies across languages have provided evidence that subject relative clauses are easier to process than object relative clauses (e.g., King and Just 1991 for English; Kwon, Polinsky, and Kluender 2006 for Korean; Lin and Bever 2006 for Chinese) and that Japanese is no exception (Kahraman 2011; Miyamoto and Nakamura 2003 from a self-paced reading study; Ueno and Garnsey 2007 from a self-paced reading study as well as an event-related potential [ERP] study).

An example of subject relatives is (1a), in which the relative head *teacher* (filler) is coindexed with the subject (gap) within the relative clause. Likewise, an example of object relatives is (1b), in which the filler *teacher* is coindexed with the gap, which is the object within the relative clause.

(1) a. *The teacher*$_i$ [*that* (gap$_i$) hated the students] quit the school.
(Subject relative)

b. *The teacher*$_i$ [*that the students hated* (gap$_i$)] *quit the school.*
(Object relative)

Explanations for this universally observed preference for subject relatives have been given from more than one perspective, e.g., language typology and markedness (Keenan and Comrie 1977), structural distance between filler and gap (O'Grady 1999, 2001), and memory load caused by linear distance between filler and gap (Gibson 2000; Wanner and Maratos 1978).

One widely cited study on L1 Japanese is Miyamoto and Nakamura (2003), which uses a self-paced reading task and compares reading times of subject relatives such as (2a) and object relatives such as (2b). Results showed that the subject relative filler *onnanoko wa* 'girl' was read significantly faster than the object relative filler 'girl', which can be taken as evidence of the processing preference for subject relatives.

(2) a. Subject relative:

| | | | | | |
|---|---|---|---|---|---|
| [*Tosiyorino* | *obaasan* | *o* | *basutei* | *made* | *miokutta*] |
| elderly | woman | ACC | bus.stop | to | accompanied |

| | | | | |
|---|---|---|---|---|
| *onnanoko* | *wa* | *nuigurumi* | *o* | *daiteita.* |
| girl | TOP | stuffed.toy | ACC | holding |

'The girl that accompanied the elderly woman to the bus stop was holding a stuffed toy.'

b. Object relative:

| | | | | | |
|---|---|---|---|---|---|
| [*Tosiyorino* | *obaasan* | *ga* | *basutei* | *made* | *miokutta*] |
| elderly | woman | NOM | bus.stop | to | accompanied |

| | | | | |
|---|---|---|---|---|
| *onnanoko* | *wa* | *nuigurumi* | *o* | *daiteita.* |
| girl | TOP | stuffed.toy | ACC | holding |

'The girl that the elderly woman accompanied to the bus stop was holding a stuffed toy.'

L2 Japanese studies on reading times also find a similar preference for subject relatives (Kahraman et al. 2009; Kahraman 2012; Kanno 2001; Kashiwagi 2011; Mitsugi, MacWhinney and Shirai 2010). For example, in Mitsugi, McWhinney and Shirai (2010), novice to lower-intermediate Korean-speaking (n = 16) and English-speaking (n = 16) Japanese as a foreign language (JFL) learners read a sentence pair such as those in (3).[1,2] Results showed that the Korean-speaking JFL learners

**1** In this chapter, for the ease of comparison of L2 learner groups across different studies, we re-define the proficiency level of the learner groups as follows. Novice learners are those who have studied Japanese less than two years in a JFL environment or those whose proficiency level is equivalent to Level 4 (N5) in *Nihongo Noryoku Shiken* or the Japanese-Language Proficiency Test. Intermediate learners are those who have studied for three to four years in a JFL environment or those whose proficiency level is equivalent to Level 3 (N3–N4). Advanced learners are those whose proficiency level is equivalent to Level 2 (N2) or above. Our definition of their proficiency levels may not be the same as the original studies.

**2** The original experiment compared three conditions: subject relatives, object relatives, and passive relatives. Given the space limitation, we only show relevant examples, and this applies to other example sentences in this chapter as well.

(but not the English-speaking learners) read the subject relative filler *kodomo ga* 'child' significantly faster than the object relative filler 'child', which provides further evidence supporting the processing ease of subject relatives.

(3) a. Subject relative:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| [*Apaato* | *de* | *yasasii* | *ruumumeeto* | *o* | *ketta*] | *kodomo* | *ga* |
| apartment | LOC | kind | roommate | ACC | kicked | child | NOM |

| | | | | |
|---|---|---|---|---|
| *kooen* | *de* | *hon* | *o* | *yonda.* |
| park | LOC | book | ACC | read |

'The child that kicked the kind roommate in the apartment read the book in the park.'

b. Object relative:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| [*Apaato* | *de* | *yasasii* | *ruumumeeto* | *ga* | *ketta*] | *kodomo* | *ga* |
| apartment | LOC | kind | roommate | NOM | kicked | child | NOM |

| | | | | |
|---|---|---|---|---|
| *kooen* | *de* | *hon* | *o* | *yonda.* |
| park | LOC | book | ACC | read |

'The child that the kind roommate kicked in the apartment read the book in the park.'

From this, Mitsugi, MacWhinney and Shirai (2010: 136) concluded that "L2 processing is not fundamentally different from L1 processing". The same preference for subject relatives is also confirmed among learners of different proficiency levels and different L1s, for example, among English-speaking JFL learners with a lower-intermediate to advanced level (Kashiwagi 2011), Turkish-speaking JFL learners with an intermediate to advanced level (Kahraman 2012), and Chinese-speaking Japanese as a second language (JSL) learners with an advanced level (Kahraman et al. 2009).

The above studies have a shared characteristic among their test sentences, namely that all of their relative clauses contain an animate subject and an animate object. For example, in (3a) above, the subject in the relative clause is the 'child', which is a gap position, and the object is the 'roommate', both of which are animate nouns. This "animate-animate" combination is especially preferred for experiments because it enables the researcher to prepare a minimal pair so that they can always compare the reading times at the identical head noun 'child' (critical region).

However, Sato (2011) argues that the corpus frequency of such relative clauses is actually very low (8.5%). Rather, he claims that the most frequent structure has an animate subject and an inanimate object in the relative clause, the "animate-inanimate" combination (75%). He further shows that native speakers of Japanese cease to exhibit the preference for subject relatives when the relative clause contains the "animate-inanimate" combination (Exp. 3).

(4) a. Subject relative with the "animate-inanimate" combination

| | | | | | |
|---|---|---|---|---|---|
| [*Kaikaku* | *o* | *katatta*] | *daitooryoo* | *wa* | *raigetu* |
| reform | ACC | mentioned | president | TOP | next month |

| | | |
|---|---|---|
| *gaiyuu* | *o* | *yotei siteiru.* |
| trip abroad | ACC | is.planning |

'The president that mentioned the reform is planning to travel abroad.'

b. Object relative with the "animate-inanimate" combination

| | | | | | |
|---|---|---|---|---|---|
| [*Daitooryoo* | *ga* | *katatta*] | *kaikaku* | *wa* | *raigetu* |
| president | NOM | mentioned | reform | TOP | next month |

| | | |
|---|---|---|
| *zissi* | *o* | *yotei siteiru.* |
| implementation | ACC | is.planning |

'The reform that the president mentioned is going to be implemented next month.'

In the sentence pair above, the object relative head 'reform' was read faster than the subject relative head 'president', though the difference was only marginally significant, which is the opposite of the results found when the relative clause contained the "animate-animate" combination. These findings suggest that the processing preference for subject relatives is subject to animacy. The preference seems to be a robust strategy under the "animate-animate" combination, but its effect is obviously reduced under the "animate-inanimate" combination. (See Kahraman and Sakai, this volume, for a detailed discussion of animacy in relative clauses.)

In the case of L2, Sawasaki (2008, 2009a) examines reading times of English-, Korean-, and Chinese-speaking learners of Japanese who were in the intermediate and advanced levels.[3] He uses the animate-inanimate combination in his test sentences such as below.

(5) a. Subject relative with the animate-inanimate combination

| | | | | | |
|---|---|---|---|---|---|
| [*Tabako* | *o* | *yameta*] | *otona* | *wa* | *yasasii.* |
| cigarette | ACC | quit | adult | TOP | kind |

'The adults who quit smoking are kind.'

b. Object relative with the animate-inanimate combination

| | | | | | | |
|---|---|---|---|---|---|---|
| [*Yooko* | *ga* | *totta*] | *syasin* | *wa* | *maamaa* | *desu.* |
| Yoko | NOM | took | picture | TOP | so so | is |

'The picture that Yoko took is so so.'

**3** Sawasaki (2008) compares intermediate JFL learners (n = 15) and upper intermediate JFL and advanced JSL learners (n = 35) whose L1 is English. In Sawasaki (2009a), the comparison is made among intermediate JFL learners whose L1 is English (n = 15) and advanced JSL learners whose L1s are English (n = 26), Korean (n = 23), and Chinese (n = 19), with the identical experimental sentences and task as in Sawasaki (2008). Some of the English-speaking participants in the two studies overlap.

Sawasaki (2008, 2009a) found that all his learner groups with different L1s and different proficiency levels as well as native speakers read the object relative head or its later region significantly faster than that of the subject relatives. These findings seem to match the results of the L1 Japanese study by Sato (2011).

The findings from Sawasaki (2008, 2009a) also found the processing preference for subject relatives at the embedded verb region 'quit/took'. The embedded verb of the object relative 'took' was read significantly more slowly than the embedded verb of the subject relative 'quit'.[4] Sawasaki (2008, 2009a) attributes these slower reading times of the object relative verb to an unnatural word sequence, that is, the sequence of the object relative (i.e., the subject and the transitive verb without the object as in 'Yoko took ø') is unnatural compared to the sequence of the subject relative (i.e., the object and the transitive verb without the subject as in 'ø quit smoking'). This unnaturalness felt at the object relative verb leads to the preference for the subject relative verb. Alternatively, however, it is also possible to claim that this difference in reading times simply reflects the difference in word recognition times between 'took' and 'quit', rather than the unnatural word sequence shown above (cf., Sawasaki (2012) for the arguments against this view). Although we need more research on this topic, if Sawasaki is correct, it means that L1 and L2 processing takes place in the same manner under the animate-inanimate combination as well.

Taking into consideration the other findings shown earlier in this section, the literature appears to point to the same conclusion, that is, L2 processing of Japanese relative clauses is identical with L1 processing regardless of the animacy of the object in the relative clause. However, a closer look at the results reveals several differences between L1 and L2 processing. First, as mentioned earlier with (3), Mitsugi, MacWhinney and Shirai (2010) found that Korean-speaking learners of Japanese showed faster reading times for the subject relatives with the animate-animate combination. In the same experiment, however, they failed to find the same effect among English-speaking learners of Japanese. The online reading times (as well as offline comprehension accuracy data measured during self-paced reading task) by the English-speaking learners did not support a processing preference for subject relatives.

Similarly, in Kashiwagi (2011), English-speaking learners (n = 14) did exhibit faster reading times for subject relatives with the animate-animate combination, but she claims that these results should be taken with caution. According to Kashiwagi (2011: 95–96), the raw reading times do not reveal any differences between the two

---

**4** This effect was absent in the L1 Japanese results by Sato (2011). More research is necessary to see why the effect was absent in Sato, but one reason could be that the vocabulary used in their test sentences was different. While Sato's sentences contained Chinese-origin words such as *kaikaku* 'reform' and *kokumu tyookan* 'secretary of the state,' there were less words of this type in Sawasaki (2008, 2009a).

types of relative clauses. The difference appeared only when the outliers (greater than two standard deviations) were replaced with the mean reading times. The subject relatives contained more outliers (greater reading times) than the object relatives, which calls into question the processing preference for subject relatives. Moreover, her offline comprehension question data showed that the accuracy rate for subject relatives was significantly lower than that for object relatives. These studies suggest that the subject relatives could be more difficult to comprehend for the learners. Both Kashiwagi (2011) and Mitsugi, MacWhinney and Shirai (2010) examined English-speaking learners with pre-advanced proficiency levels. Possible reasons for the absence of the subject relative preference could be that their L1, English, is typologically different from Japanese and thus the learners were confused, and/or that their low proficiency of Japanese inhibited native-like processing.

In addition to Mitsugi, MacWhinney and Shirai (2010) and Kashiwagi (2011), the absence of the processing preference for the subject relatives was also observed in Kanno (2007), which employed an offline picture selection comprehension task. The participants listened to the example sentence twice and selected the matching picture out of three candidates. Kanno (2007) compared subject relatives and object relatives with the animate-animate combination and also with the animate-inanimate combination. The participants were at the novice level with different L1s: Chinese (n = 43), Vietnamese (n = 10), Indonesian (n = 10), Thai (n = 7), and Sinhalese (n = 8). The results were inconclusive. First, under the animate-animate combination (when the subject NP and object NP could be semantically and pragmatically reversed),[5] the comprehension accuracy rate was very low, and no group showed significantly higher accuracy for the subject relatives. However, when the relative clauses with the animate-inanimate combination were compared (when the subject NP and object NP could not be semantically and pragmatically reversed), the comprehension accuracy was significantly higher for the subject relatives among Chinese, Vietnamese, and Indonesian speakers. Kanno (2007: 214) concludes that a semantic cue such as animacy is useful for her participants because novice learners may find it difficult to utilize case information appropriately. Moreover, she claims that when the sentences were perceived as difficult, their interpretation was influenced by L1 features such as word order (see also Yamashita (2008) for a similar discussion).

**5** While Sato (2011) relies on the corpus frequency to explain the processing difficulties between the animate-animate condition and the animate-inanimate condition, Kanno (2007) explains it in terms of the reversibility of the subject NP and the object NP.

**Table 1:** Summary of the previous findings for the animate-animate combination

| Studies | L1 | L2 Proficiency | Method (what is compared) |
|---|---|---|---|
| **Evidence indicating L1 and L2 similarities (preference for subject relatives):** | | | |
| Kahraman et al. (2009) | Chinese (n = 19) | Advanced | Self-paced (reading times for whole sentence) |
| Kahraman (2012) | Turkish (n = 26) | Intermediate to advanced | Self-paced (reading times per region) |
| Mitsugi, MacWhinney, and Shirai (2010) | Korean (n = 16) | Lower intermediate | Self-paced (residual reading times per region) |
| Kanno (2001) | English (n = 17) | Intermediate | Self-paced (reading times for whole sentence) |
| Kashiwagi (2011) | English (n = 14) | Lower intermediate to advanced | Self-paced (residual reading times per region) |
| **Evidence indicating L1 and L2 differences:** | | | |
| Mitsugi, MacWhinney, and Shirai (2010) | English (n = 16) | Lower intermediate | Self-paced (residual reading times per region) |
| Kashiwagi (2011) | English (n = 14) | Lower intermediate to advanced | Self-paced (raw reading times per region) |
| Kanno (2007) | Chinese (n = 43), Vietnamese (n = 10), Indonesian (n = 10), Thai (n = 7), Sinhalese (n = 8) | Novice | Picture selection (multiple choice) |

In sum, the previous findings of L2 Japanese processing for relative clauses are inconclusive. Tables 1 and 2 show a summary of the findings. Although it seems as though L2 learners can generally process subject relatives and object relatives in a native-like manner, there still are several differences between L1 and L2 processing that cannot be ignored, especially when the learners' L1 is English or when they are novice learners. The same conclusion is true for additional studies that could not be mentioned in this section because of space limitation, for example, Roberts (2000), which analyzed offline comprehension and production data; Yamashita (2008), which examined offline comprehension and translation data; Currah (2004), which looked at online and offline data of gapless relative clauses (NP modifying clause with no coindexed gap). These studies all included English-speaking JFL learners as their participants and most of them were at the pre-advanced level. Their results also found that the learners processed relative clauses in a somewhat different manner from native speakers.

**Table 2:** Summary of the previous findings for the animate-inanimate combination

| Evidence indicating L1 and L2 similarities (non-preference for subject relatives): | | | |
|---|---|---|---|
| Studies | L1 | L2 Proficiency | Method (what is compared) |
| Sawasaki (2009a) | English (n = 41), Korean (n = 23), Chinese (n = 19) | Intermediate (English, n = 15) Advanced (English, Korean, Chinese) | Self-paced (reading times per region per mora) |
| Sawasaki (2008) | English (n = 50) | Intermediate (n = 15) and upper intermediate to advanced (n=35) | Self-paced (reading times per region per mora) |
| Kanno (2007) | Thai (n = 7) Sinhalese (n = 8) | Novice | Picture selection (multiple choice) |
| Evidence indicating L1 and L2 differences: | | | |
| Studies | L1 | L2 Proficiency | Method (what is compared) |
| Kanno (2007) | Chinese (n = 43), Vietnamese (n = 10), Indonesian (n = 10) | Novice | Picture selection (multiple choice) |

Such nonnative-like processing behavior of the learners could be due to their L2 development, suggesting that they are still on the way to acquiring native-like processing. The nonnative-like behavior could also be related to L1 influence, suggesting that typological differences between Japanese and English are causing processing difficulties. Whichever is the case, a subsequent question emerges – can all learners despite their L1 ultimately attain native-like processing, or is native-like processing difficult for certain groups of L1 speakers even after reaching the advanced level? Findings in Sawasaki (2008, 2009a), which examined English-speaking learners, appear to support the former, but we still need to wait for future research to fully answer these questions as his studies deal only with the animate-inanimate combination of the relative clauses.

Apart from the relative clause construction, there is some evidence showing that L2 processing continues to struggle to become native-like even after the learner's proficiency reaches the advanced level. In the next section, we will see that the processing of a very simple sentence can highlight processing differences between L1 and L2 even after the L2 level becomes proficient.

## 2.2 Similarities and differences in incremental processing

Previous studies provide ample evidence showing that L1 readers do not wait to process until the end of the sentence. Rather, processing happens as a sentence unfolds (incremental processing), and this is true in English (Altmann and Kamide 1999; Frazier and Rayner 1982; Trueswell, Tanenhouse and Kello 1993; etc.) and in

typologically different languages such as Japanese (Kamide, Altmann and Haywood 2003; Miyamoto 2002; Yamashita 1994, 1997; etc.). Regardless of the language, previous research shows that readers make use of available information that is helpful for incremental processing, for example, verb information in English and case information in Japanese.

Only a few studies, however, have been reported on L2 Japanese incremental processing (e.g., Lieberman, Aoshima and Phillips 2006; Mitsugi 2011; Sawasaki 2004, 2007; Zhai 2009). Among these, Lieberman, Aoshima and Phillips (2006) examined whether native speakers of Japanese (n = 24) and English-speaking learners of Japanese (n = 15) at an advanced level would complete the following fragments as direct questions or indirect questions.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (6) | a. | *Sensei* | *wa* | *seito* | *ga* | *tosyositu* | *de* | *dare* | *ni…* |
| | | teacher | TOP | student | NOM | library | LOC | who | DAT |
| | b. | *Dare* | *ga* | *sensei* | *ni* | *seito* | *ga* | *tosyositu* | *de…* |
| | | who | NOM | teacher | DAT | student | NOM | library | LOC |

The above fragments all contain a *wh*-phrase *dare* ‘who’ either at the sentence-initial nominative position or at the sentence-internal dative position. Japanese is a *wh*-in-situ language, thus (6a) allows both direct and indirect questions whereas (6b) only allows a direct question. An example of the direct question is (7a), where the question marker (QM) *ka* appears at the sentence-final position, and an example of the indirect question is (7b), where the QM appears at the end of the embedded clause. This is because of the Japanese requirement that “the QM must be at least as high in the sentence structure as the thematic position of the *wh*-phrase” (Lierberman, Aoshima and Phillips 2006: 426).

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| (7) | a. | *Sensei* | *wa* | *seito* | *ga* | *tosyositu* | *de* | *dare* | *ni* | *au* | *to* |
| | | teacher | TOP | student | NOM | library | LOC | who | DAT | see | COMP |
| | | *omoimasita* | *ka.* | | | | | | | | |
| | | thought | Q | | | | | | | | |

‘Who did the teacher think that the student saw in the library?’

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | b. | *Sensei* | *wa* | *seito* | *ga* | *tosyositu* | *de* | *dare* | *ni* | *au* | *ka* |
| | | teacher | TOP | student | NOM | library | LOC | who | DAT | see | Q |
| | | *sirimasen.* | | | | | | | | | |
| | | know.not | | | | | | | | | |

‘The teacher doesn’t know who the student will see in the library.’

On the other hand, English does not allow a *wh*-in-situ construction. Colloquial English occasionally allows an echo question such as *You said that John saw who?* but it is interpreted only as a direct question.

The results of the experiment show that a vast majority of both native speakers and Japanese learners completed (6a) as an indirect question and (6b) as a direct question. The authors claim that both L1 and L2 participants are guided by the same processing mechanism when resolving the *wh*-scope ambiguity, trying to anticipate the earliest available position for the QM, thus producing indirect questions for (6a). Though these findings emerge from a production study rather than a processing study, these results are still convincing evidence of incremental processing in L2 Japanese.

In Mitsugi (2011), lower-intermediate learners of Japanese (n = 38, a mixed L1 group of Korean, Chinese, and English) read sentences such as those below. Sentence (8a) is ungrammatical because Japanese does not allow two accusative NPs appearing in the same clause, according to the so-called "double-o constraint".

(8) a. **Ohuisu de isogasii syain ga kibisii syatyoo o*
office LOC busy employee NOM strict president ACC
*iyaiya atui otya o dasita.*
unwillingly hot tea ACC served

b. *Ohuisu de isogasii syain ga kibisii syatyoo ni*
Office LOC busy employee NOM strict president DAT
*iyaiya atui otya o dasita.*
unwillingly hot tea ACC served
'In the office, the busy employee unwillingly served the strict president the hot tea.'

c. *Ohuisu de isogasii syain ga kibisii syatyoo o*
office LOC busy employee NOM strict president ACC
*iyaiya karakatta.*
unwillingly teased
'In the office, the busy employee unwillingly teased the strict president.'

When reading times were compared, both native speakers and L2 learners exhibited longer reading times at the second accusative NP 'hot tea' in (8a) than at the first accusative NP 'hot tea' in (8b). The longer reading times suggest that the ungrammaticality of sentence (8a) was detected at the second accusative NP, the very place where the double-o constraint is violated, rather than waiting until the end of the sentence.

Next, Mitsugi (2011) compared sentences (8b) and (8c) using the visual-world paradigm task. The visual-world paradigm task typically asks participants to look at pictures or objects while listening to a sentence, and their eye-movements at critical region(s) are recorded (cf., Kamide, Altmann and Haywood 2003). The participants

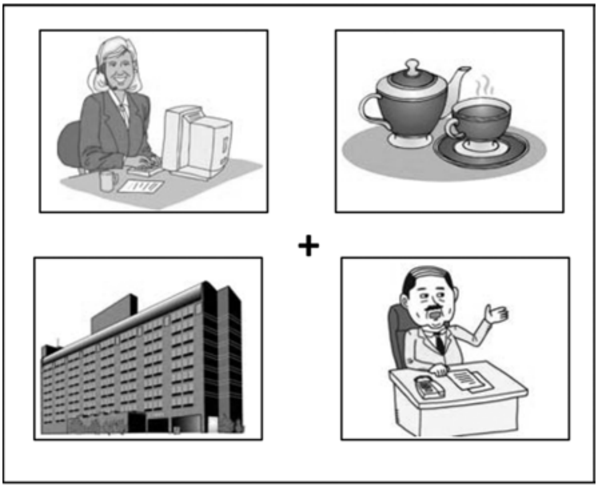


**Figure 1:** Visual stimuli for the visual-world paradigm task (Mitsugi 2011: 119)

were presented with four different pictures such as the ones shown in Figure 1 while listening to (8b) or (8c). The aim was to examine which picture the participants would look at when they heard the word *iyaiya* 'unwillingly'. If the participants perform incremental processing utilizing available information, they would gaze at the picture of the tea in the cup at 'unwillingly' in (8b). This result is predicted because the combination of the nominative-marked 'employee' (typically an agent) and the dative-marked 'president' (typically a recipient) would allow them to anticipate an accusative NP, something transferrable (tea in the cup), to follow. However, the same prediction would not hold in the case of (8c), where the combination of the nominative-marked 'employee' (agent) and the accusative-marked 'president' (theme) would not necessarily require another NP to follow. This prediction was borne out. Both native and learner groups cast more frequent and faster anticipatory looks toward the tea in the cup at 'unwillingly' in (8b) than in (8c). These findings suggest that not only can L2 learners start to process sentences early on, but they can also anticipate what word would appear next, in the same way as native speakers do.

Thus far, we have seen findings indicating that L1 processing and L2 processing are similar in terms of incremental processing. However, there is some evidence showing that L1 and L2 are not always processed in the same manner.

Sawasaki (2004, 2007) examined the incremental nature of L2 Japanese, using short, simplex sentences with no garden-path or ambiguity involved.[6] His findings

**6** In a garden-path sentence, one is first led to a misanalysis and then forced to revise the structure and meaning accordingly. A well-known garden-path example is *The horse raced past the barn fell*. Initially, this sentence is most likely interpreted as 'The horse raced past the barn', but it needs to be reanalyzed as 'The horse that was raced past the barn fell'.

indicate that L2 Japanese learners perform incremental processing but in a different way from native speakers. For example, the participants in Sawasaki (2007) were Japanese learners who were native speakers of English (intermediate, n = 15, upper intermediate, n = 9, and advanced, n = 26), Korean (advanced, n = 23), and Chinese (advanced, n = 19). Test sentences were as follows.[7]

(9) a. *Tosyokan* *de* *Tanaka-san* *ga* *bikkurisita* *to*
library LOC Tanaka NOM was.surprised COMP
*iimasita*. (Int1 type)
said
'Tanaka said that s/he was surprised at the library.'

b. *Tosyokan* *de* *Tanaka-san* *ga* *kinoo* *bikkurisita* *to*
library LOC Tanaka NOM yesterday was.surprised COMP
*iimasita*. (Int2 type)
said
'Tanaka said that s/he was surprised at the library yesterday.'

c. *Tosyokan* *de* *Tanaka-san* *ga* *kinoo* *totemo* *bikkurisita*
library LOC Tanaka NOM yesterday very was.surprised
*to* *iimasita*. (Int3 type)
COMP said
'Tanaka said that s/he was very surprised at the library yesterday.'

d. *Tosyokan* *de* *Tanaka-san* *ga* *sensei* *o* *tetudatta* *to*
library LOC Tanaka NOM teacher ACC helped COMP
*iimasita*. (Tran2 type)
said
'Tanaka said that s/he helped the teacher at the library.'

e. *Tosyokan* *de* *Tanaka-san* *ga* *kinoo* *sensei* *o*
library LOC Tanaka NOM yesterday teacher ACC
*tetudatta* *to* *iimasita*. (Tran3 type)
helped COMP said
'Tanaka said that s/he helped the teacher at the library yesterday.'

Using these sentences, Sawasaki (2007) examined if the learners differentiate arguments (accusative NPs such as 'teacher') and adjuncts (time/degree adverbs

7 Specifically, the test sentences are biclausal and not simplex. However, the regions of interest are regions before reaching the embedded verb "was surprised". Up until this region, we assume the test sentence is interpreted as a "simplex sentence" because it is considered the most cost-free way of processing.

such as ‘yesterday’ and ‘very’) before reaching a verb. Previous studies in L1 English processing show that the argument word is read faster than the adjunct word (Ahrens 2003; Boland and Blodgett 2006; Frazier and Clifton 1996; etc.). In the above sentence set, reading times of ‘teacher’ in (9d) and ‘yesterday’ in (9b), (9c), and (9e) were compared in the region of the third content word. Additionally, reading times for ‘teacher’ in (9e) and ‘very’ in (9c) were compared in the region of the fourth content word. If arguments are also read faster in an SOV language like Japanese, it would mean arguments are distinguished from adjuncts pre-verbally, constituting evidence for incremental processing.

Results showed that native speakers of Japanese read ‘teacher’ significantly faster than ‘yesterday’ in the third region, but the same participants read ‘teacher’ in the fourth region significantly more slowly than ‘very’. Sawasaki (2007) interprets these results to mean that arguments are not necessarily read faster than adjuncts, but rather, a word enabling the reader to predict the following word leads to faster reading times. For example, a degree adverb ‘very’ typically modifies a subsequent predicate, and thus the reader can anticipate a certain group of words to follow (e.g., a verb or adjective such as *very surprised*). On the other hand, a time adverb ‘yesterday’ is typically a sentence modifier, and thus it is not as easy to predict the subsequent word compared to a degree adverb (e.g., *surprised yesterday* is not as predictable a modification as *very surprised*). Native speakers in the experiment were sensitive to this difference between degree adverbs and time adverbs, and it lead to faster reading times of the degree adverb than the argument NP on the one hand, and slower reading times of the time adverb than the argument NP on the other hand.

Similar, though a little weaker, results emerged from advanced learners with L1 English and L1 Korean as well, indicating that they were also sensitive to the degree vs. time adverb differences during online processing. However, advanced learners with L1 Chinese and intermediate learners with L1 English failed to exhibit the same sensitivity. The results of the advanced Chinese-speaking learners were unexpected, but these results were interpreted in terms of the experimental artifact. That is, reading times were adjusted and compared in terms of residual reading times per “mora”, which presumably obscured the reading times of Chinese speakers who may tend to read Chinese characters with “character” or “syllable” as the base unit instead of mora as the base unit.[8]

---

**8** In this study, residual reading times (RRTs) are obtained on the basis of expected reading times (cf. Mazuka, Itoh and Kondo 2002; Trueswell, Tanehous and Garnsey 1994). RRTs are calculated for each participant by using the formulas below, where “a” and “b” are parameters varying among participants.

(i) a. Expected reading times = a + b*(number of moras)

b. RRTs = Raw reading times − Expected reading times

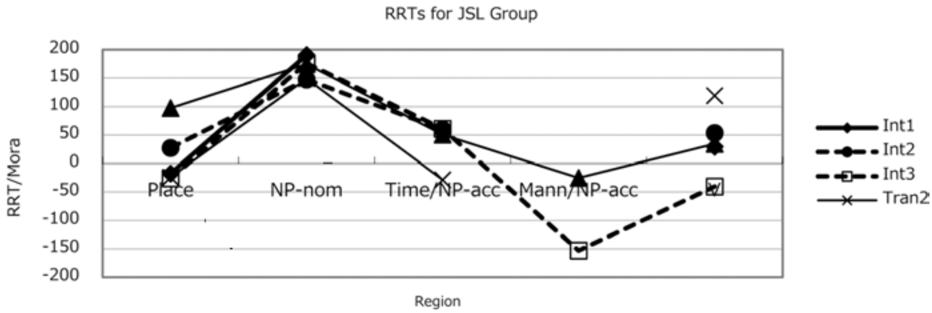



**Figure 2:** Reading times of advanced learners with L1 English (Sawasaki 2007: 72)[9,10]

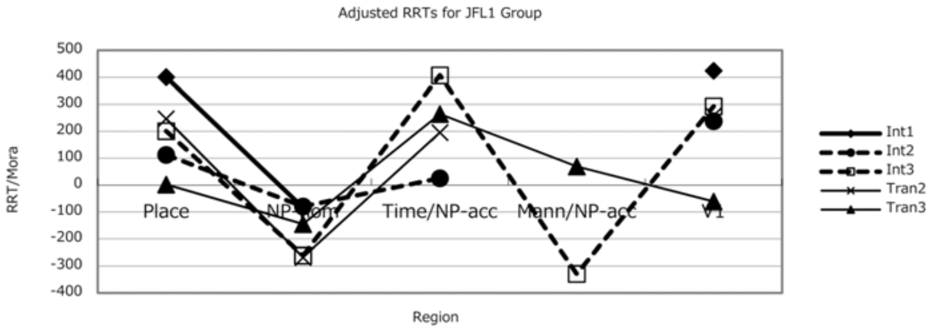



**Figure 3:** Reading times of intermediate learners with L1 English (Sawasaki 2007: 80)

Intermediate learners with L1 English showed a very different reading pattern from advanced learners with L1 English. For advanced learners, the reading time peaks were found at the nominative NP region 'Tanaka' and the first verb region 'was surprised' (Figure 2), which can be interpreted to mean that they distributed special attention to the subject and verb information to mark a clause boundary. However, for intermediate learners, the reading time peaks were found at every other region (Figure 3), which can be interpreted to mean that they stopped to interpret the sentence meaning every two words. From these findings, Sawasaki (2007) claims that advanced learners and intermediate learners perform different processing strategies, but both groups share a common strategy whereby they start to process a sentence incrementally at an early region in ways that best suit their proficiency levels.

9 Figures 2 to 4 exclude reading times for the matrix verb, the sentence final region.
10 Positive RRTs indicate slower reading times and negative RRTs indicate faster reading times (see note 8).

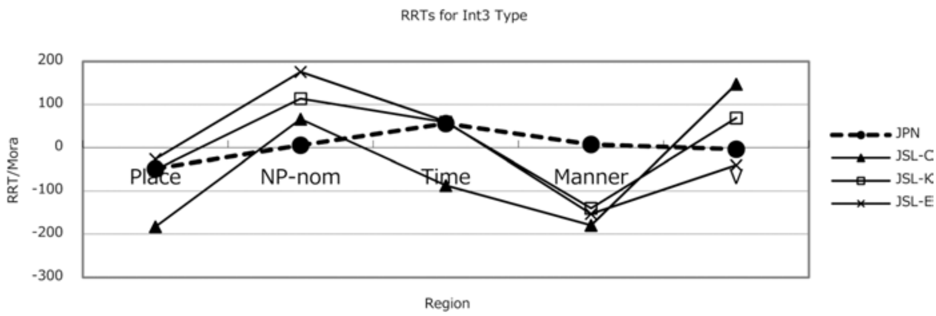



**Figure 4:** Reading times of the Int3 sentence by native speakers and advanced learners with L1 Chinese (JSL-C)/Korean (JSL-K)/English (JSL-E) (Sawasaki 2007: 158)

Furthermore, Sawasaki (2007) also found that the reading patterns of advanced learners (L1 Chinese, Korean, and English) share some commonalities that were not observed among native speakers of Japanese. For example, Figure 4 compares reading times of the Int3 sentence, (9c), and the results clearly demonstrate that all three advanced learners experience reading time peaks at the nominative region and the clause-final verb region, as discussed earlier. However, native speakers do not exhibit the same reading time peaks. Interestingly, a similar pattern emerged in other sentence conditions as well. This finding suggests that although advanced learners process time/degree adverbs in a similar manner as native speakers, it does not mean they perform native-like processing in all respects (cf. Clahsen 2011; Clahsen and Felser 2006 for the view that L2 processing cannot reach native-like processing).

Sawasaki (2007) explains that these differences are caused by L2 cognitive restrictions that result from learners' lesser amounts of exposure to Japanese compared to native speakers. Sawasaki (2007: 181) further argues that "learners may not have to perform exactly native-like processing unless failure to do so causes a serious problem." Since current test sentences are very short, simple sentences with no garden-path or ambiguity involved, processing in learners' own ways, which are different from native speakers, will not immediately harm successful comprehension of the sentence. In a case like this, differences between L1 and L2 processing may persist even after the learners reach the advanced level. However, the more complicated and/or longer the sentences become, the more necessary it may become that the learners adjust their reading strategies to match native speakers in order to perform successful and efficient processing.

This section has discussed similarities and differences between L1 and L2 Japanese processing by focusing on relative clauses and incremental processing. Many previous studies have shown that L1 and L2 processing strategies are very similar in that subject relatives are generally easier to process under the animate-animate combination and

that all readers try to perform incremental processing. However, some studies suggest other findings. These discrepancies between studies may stem from experimental artifacts, developmental reasons, L1 influence, L2 cognitive restriction, and/or other factors that are beyond the scope of this chapter. Whether these discrepancies are trivial or fundamental to L2 processing is unclear at this moment, but taking a closer look at them is important because they move us one step closer to capturing a more accurate picture of L2 processing.

# 3 Individual differences: Influence of working memory on L2 Japanese sentences processing

Many factors contribute to the differences in L2 learners' sentence processing performance. For instance, factors that affect L2 sentence processing include L2 learners' L1, age of when L2 learning started, overall L2 proficiency level, learners' cognitive skills, and learners' working memory capacity. In this section, we will specifically discuss L2 sentence processing and its relation to working memory capacity, which is one of the cognitive processes thought to have an impact on sentence processing and comprehension. Particularly, we will review recent findings in the literature about L2 Japanese sentence processing and working memory.

## 3.1 Working memory and sentence processing

Acquiring a second/foreign language and using it requires a great deal of cognition. Among the many different cognitive processes required for comprehending a language, working memory is said to play an extremely important role. The notion of working memory, which was made popular by Baddeley and Hitch (1974), refers to the ability to maintain information while manipulating and integrating other information required for the task at hand (Baddeley 2003; Baddeley and Logie 1999). Many studies have been carried out to examine the effect of working memory on sentence processing, especially in L1 because successful sentence processing and comprehension require one to process the incoming elements in addition to storing, maintaining, and integrating other pieces of information such as lexical, syntactic and discourse information.

Baddeley and Hitch's original model of working memory had three components: a control system called the central executive, and two storage systems, namely the visuo-spatial sketchpad and the phonological loop. The phonological loop stores memory traces for a few seconds before it fades and it will also carry out articulatory rehearsal, which is the process of subvocally repeating material to prevent from forgetting. The visuo-spatial sketchpad can hold and manipulate visual and spatial

representations. The central executive was originally characterized as a pool of general processing capacity that dealt with issues that were not taken care of by the two storage systems. Baddeley and Hitch's original three-component working memory model encountered problems when the interactions between long-term memory and working memory were considered. To account for the shortage in the original model, a fourth component, the episodic buffer, was proposed. The function of the episodic buffer is to bind and retrieve information to form integrated episodes (Baddeley 2003).[11] Among these components, the functions of the phonological loop and the central executive control are argued to be extremely important in L1 and L2 acquisition, learning, and use (Juffs and Harrington 2011).

Different working memory models have been proposed. For example, a "capacity-constrained" model proposed by Just and Carpenter (1992) suggests that one single working memory system is shared among memory (storage) and computational operations (efficiency) and there are tradeoffs between storage and processing. Just and Carpenter's model suggests that an individual has a different amount of resources for processing and storing and that individual differences in performance are caused by the amount of resources that one has, that is, capacity. Another model proposed by Caplan and Waters (1996, 2002) states that sentence processing is mediated by two separate working memory systems used for online processing (interpretive processing) and offline processing (post-interpretive processing). In their model, syntactic processing is part of the interpretive processes and individual differences occur because of domain-specific working memory capacity. Another model by MacDonald and Christiansen (2002) emerges from the connectionist framework and assumes that an individual's experience in language (capacity) and phonological representations (efficiency) affects their processing abilities. Working memory in their model is considered the "network itself" (MacDonald and Christiansen 2002: 38) and the individual differences found in processing abilities emerge from an individual's unique experience with language and biological differences such as phonological representations developed for language. Despite the differences among these various models, they all seem to agree on the effect of working memory on sentence processing.

The size of working memory varies from individual to individual (Daneman and Carpenter 1980; Just and Carpenter 1992; Osaka and Osaka 1992), and these differences affect the way that sentence processing is carried out (Nakano, Felser and Clahsen 2002; Sawasaki 2009b; Tokimoto 2004). In order to study the effect of working memory capacity on sentence processing, two approaches have mainly been taken. The first is to divide participants into high and low working memory capacity groups, often by using a version of the Reading Span Test (Daneman and Carpenter 1980, commonly referred to as "RST"), and compare group differences in sentence processing and comprehension (e.g., King and Just 1991). The second is the dual-task approach. This approach asks participants to do a secondary task, which is

11 For other proposals on working memory structure, see for example, McElree (2006).

usually manipulated in terms of the load or number of items held in memory, while performing sentence comprehension. This approach is based on the assumption that if primary and secondary tasks rely on the same working memory resources, then having an increased load on memory with the secondary task will decrease the performance of sentence comprehension (e.g., Fedorenko, Gibson and Rohde 2004).[12]

Many studies that investigated the relationship between working memory capacity and L1 sentence processing have used Daneman and Carpenter's (or a similar style) reading span test, which is argued to tax both storage and processing resources.[13] The procedure for administering their reading span test is as follows. The researcher shows a participant a set of sentences written on 5" by 7" index cards, one sentence per card. The task for the participant is to read the sentences aloud and memorize the target word, which is located at the end of the sentence and underlined in red. After reading the sentences in each set, a white index card is presented to the participant to recall the target words from each sentence in the set. The number of sentences in each set steadily increases (from a two-sentence to a five-sentence condition), and there are five trials for each sentence condition. Although this measurement has been widely used in the field, Daneman and Carpenter's reading span test and its variations have faced some criticisms (see Roberts and Gibson 2002; Waters and Caplan 1996). Such criticisms include questions like whether the reading span test is really taxing both storage and processing resources and/or whether participants are actually processing the sentences for comprehension and not merely concentrating on memorizing the target word. Despite these criticisms, the test has been extensively used in a large number of previous studies as a working memory measurement for language processing and its predictive power of sentence comprehension has been investigated and confirmed (see Daneman and Merikle (1996) for a review of 77 previous studies, and Conway et al. (2005) for the reliability and validity of the task).

In general, participants with higher working memory capacity perform better in terms of comprehension accuracy and faster reading time.[14] For instance, King and Just (1991) used center-embedded subject and object relative clause sentences with a self-paced moving window technique in English. They divided their participants into high and low working memory span groups and their results showed that the verbs of object relative clauses in sentences like the following require more time to read for the low working memory span group than the high working memory span group.

**12** There have been also studies that used participants with selective impairments in either memory or syntactic comprehension as the participants of a study instead of normal (non-impaired) participants. Along with a working memory measurement such as a reading span test or a dual-task, the studies taking this approach examine how participants with selective impairments perform sentence processing and comprehension (e.g., Caplan and Waters 1995).

**13** Daneman and Carpenter (1980) also created the Listening Span Test.

**14** See Miyake and Shah (1999) for the limitation of working memory.

(10) *The reporter that the senator attacked admitted the error.*

They also found that the lower working memory span group's comprehension accuracy was lower than that of the higher working memory span group. Additional studies have been conducted, especially in English, to provide evidence of the relationship between working memory capacity and individual differences in sentence processing and comprehension (MacDonald, Just and Carpenter 1992; Miyake, Carpenter and Just 1994; Waters and Caplan 1992; for summaries of relevant literature see Chipere 2003; Daneman and Merikle 1992; Friedman and Miyake 2004; MacDonald and Christiansen 2002).

In L1 Japanese, the reading span test by Osaka and Osaka (1992, 1994) has been widely used to measure working memory capacity.[15] Osaka and Osaka's Japanese reading span test is slightly different from that of Daneman and Carpenter's because of the nature of Japanese language. Since many Japanese sentences end with a verb, in order to have various categories of words as target words in their Japanese reading span test, the target word was not the last word of the sentence. Using Osaka and Osaka's Japanese reading span test, previous studies in L1 Japanese have also shown that the difference in working memory capacity influences Japanese sentence as well as text comprehension (Jincho and Mazuka 2011; Kashiwagi 2011; Nakano, Felser and Clahsen 2002; Osaka and Osaka 2002; Sawasaki 2009b; Tokimoto 2004).

The topic of whether there is a correlation between L1 and L2 working memory capacities has also been examined. Previous studies show that the relationships between L1 and L2 working memory are not language-specific (Juffs 2005; Osaka and Osaka 1992; Osaka, Osaka and Groner 1993). Correlations have been found between L1 reading span test scores and L2 reading span test scores. That is, if one has high working memory capacity in L1, then that individual tends to also have high working memory capacity in L2 as well (Harrington and Sawyer 1992; Ikeno 2006; Juffs 2004, 2005; Miyake and Friedman 1998; Osaka and Osaka 1992; Osaka, Osaka and Groner 1993), and this seems to be the case not only with two languages but also in L1, L2, and L3 (Van den Noort, Bosch, Haverkort and Hugdahl 2006). Despite the correlations between L1 and L2 working memory capacity, only L2 working memory capacity seems to be a good indicator of successful L2 sentence comprehension and the effect of L1 working memory capacity on L2 sentence comprehension seems to be indirect (Alptekin and Erçetin 2010; Chun and Payne 2004; Miyake and Friedman 1998; Walter 2004).

**15** There are also RSTs created by Itomitsu and Nakayama (2005) and Watanabe (2012) for Japanese as L2 learners. Itomitsu and Nakayama's reading span test has 70 sentences that use the vocabulary (except for a few katakana words) and structural patterns from Levels 3 and 4 of the Japanese Language Proficiency Test, in an attempt to control the familiarity of vocabulary and structural patterns. Watanabe's reading span test also has 70 sentences selected from textbooks from Japanese elementary school. The reading span test sentences in Watanabe were modified in order to maintain similar sentence length with each other and some unfamiliar words were replaced with easier words as well as furigana glosses over the kanji characters.

## 3.2 L2 sentence processing and working memory

The L2 learners with high working memory capacity, in general, are assumed to be able to utilize L2 linguistic information more efficiently than learners with low working memory capacity, which may allow learners to achieve higher proficiency and/or facilitate learning (Daneman and Green 1986; Fortkamp 1999; Harrington and Saywer 1992; Miyake and Friedman 1998). Despite this fact, studies of working memory capacity and sentence processing in L2 English and non-Japanese languages show mixed results as to whether working memory capacity has any effect on the learners' L2 sentence processing and comprehension (for an overview of the effect of working memory on L2, see Juffs and Harrington 2011). Juffs (2004), for example, re-analyzed the data from Juffs (2000, 2002) with four different L1 groups of participants to investigate L2 English sentence processing. The purpose of the study was to examine whether the individual differences in working memory capacity would explain the individual variation in the learners' L2 sentence processing performance. Three types of sentences, with and without ambiguity, and garden-path sentences were used as the test sentences. Additionally, the participants were given sections of the Michigan Test of English as a second language (ESL), the reading span test in L2 (English) and their L1, and two word-span tests, a task in which participants are presented with a series of words and have to recall them, in their L1 and L2 (English).

The results in Juffs's study showed no correlations between any working memory span measures, including the L1 and L2 reading span tests, and the reading time of the region in garden path sentences, where processing load is assumed to be the greatest. However, when word span was considered, lower span learners took longer to process all the test sentences than the higher span learners. Juffs interpreted these results as evidence for a weak relationship between working memory capacity and L2 sentence processing (cf. Felser and Roberts 2007; Juffs 2005; Omaki 2005).

In contrast, other studies have shown a clear effect of working memory capacity on L2 sentence processing. For example, Sagarra and Herschensohn (2010) tested beginning and intermediate adult English-speaking learners of Spanish and Spanish monolinguals using an online self-paced reading task and an offline grammaticality judgment task containing noun-adjective sentences varied in terms of gender, number, agreement and disagreement in addition to a working memory test. They found that all groups were accurate in the judgment task, but only the intermediate learners and Spanish monolinguals exhibited sensitivity to gender and number violations in the online task. They also found that the intermediate learner group with high working memory capacity was more accurate on comprehension questions than the intermediate learner group with low working memory capacity. Their results indicate that both proficiency level and working memory capacity may affect L2 online sentence processing (cf. Dussias and Piñar 2010; Keating 2010; Ren 2009; Rodriguez 2008). Additionally, a study by Havik, Roberts, van Hout, Schreuder and Haverkort (2009),

which tested Dutch L2 subject-object ambiguous sentences, adds an interesting finding to this topic. These researchers found that there was an effect of working memory capacity on reading comprehension and parsing decisions, but the L2 high working memory group's sentence processing was patterned like native speakers with low working memory capacity. This result indicates that L2 learners with high working memory capacity may be processing sentences in a native-like manner, which may be evidence for one of the fundamental questions that the field of L2 sentence processing investigates.

A fair number of studies have been conducted on non-Japanese languages and the effect of working memory capacity on L2 sentence processing, but there has not been enough research on L2 learners of Japanese. To our knowledge, among the studies published up to 2011, only three studies have directly addressed this issue: two studies by Fukuda (2004, 2005) that investigated Malaysian native L2 learners of Japanese (2004), another that investigated Chinese native L2 learners of Japanese (2005), and Kashiwagi (2011), which investigated the effect of working memory capacity on L2 sentence processing by English native L2 learners of Japanese.

Fukuda's two studies focused on the effect of working memory capacity on listening comprehension accuracy rates and not necessarily L2 Japanese online sentence processing. For example, Fukuda (2004) tested whether working memory capacity or short-term memory span has an effect on listening comprehension and whether that effect is similar among different L2 proficiency level groups by using a listening comprehension test, a digit span test, and a listening span test. The digit span test is another type of working memory span task similar to the word span test. It uses a series of numbers as the memorization target and asks participants to immediately repeat them back. Fukuda (2004) tested Malaysian native L2 learners of Japanese whose proficiency levels were either *Japanese Language Proficiency Test* (JLPT) Levels 2 or 3, and found correlation between the listening comprehension accuracy and working memory capacity in the lower proficiency level group. However, no correlation was found in the higher proficiency level group. This result was similar to Osaka (2002) who tested Japanese native learners of English and Italian, and found stronger correlations between the learners with a shorter length of study and the working memory capacity than the learners with a longer length of study.

Assuming that the listening mechanism of higher proficiency level learners would be similar, if not the same, to that of the native speakers, Fukuda (2005) predicted that there would be a correlation between the listening comprehension accuracy and the working memory capacity, similar to the result obtained in the L1 study by Daneman and Carpenter (1980). Chinese native speakers learning Japanese with JLPT Levels 1 and 2 participated in this study, and she found no correlation between the reading span test, digit span test, and L2 listening comprehension accuracy in either Levels 1 or 2. In addition, there was no difference between the reading span test and digit span test between these groups. The findings of this study are interesting because no correlations were found between the working memory capacity and

the listening comprehension ability despite her prediction based on Fukuda (2004). This discrepancy in the results of Fukuda (2004) and Fukuda (2005) clearly suggest that more research is necessary on this issue.

The study by Kashiwagi (2011) examined the relationship between individual differences in one's working memory capacity and L2 Japanese sentence processing using the following relative clause sentence structures.

(11) a. Japanese subject-gap relative clause[16]

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| [*Ueno-san* | *o* | *miokutta*] | *ozisan* | *ga* | *omotya* | *o* | *hirotta.* |
| Mr. Ueno | ACC | saw off | uncle | NOM | toy | ACC | picked up |

'The uncle who saw Ueno off picked up a toy.'

b. Japanese object-gap relative clause

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| [*Ueno-san* | *ga* | *miokutta*] | *ozisan* | *ga* | *omotya* | *o* | *hirotta.* |
| Mr. Ueno | NOM | saw off | uncle | NOM | toy | ACC | picked up |

'The uncle who Ueno saw off picked up a toy.'

Ten English native learners of Japanese, who were divided into two working memory capacity groups – L2 RST High and L2 RST Low – using the L2 Japanese reading span test created by Itomitsu and Nakayama (2005), participated in the Japanese online sentence reading experiment.[17] The test sentences were divided into regions and presented region-by-region using a self-paced moving window task. The interesting finding from Kashiwagi's study is that the high and low working memory capacity groups, measured by the L2 Japanese reading span test, demonstrated different reading patterns.

Figure 5 illustrates that the L2 RST High group demonstrates similar reading patterns in both the Japanese subject-gap and object-gap relative clause sentences. That is, the L2 RST High group displayed a longer reading time in Region 2 (*miokutta*) and a shorter time in Region 3 (*ozisan ga*) then once again a longer time in Region 4 (*omotya o*), making a zigzag reading pattern for both sentence types. On the other hand, the L2 RST Low group shows different reading patterns for the two gap type sentences.

The analysis of the residual reading times of the different regions revealed that there were differences between the L2 RST High and Low groups in the two relative clause type sentences. First, in the Japanese subject-gap sentences, the difference

**16** In this chapter, we will use the words subject relatives and subject-gap relative clause as well as object relatives and object-gap relative clause interchangeably.

**17** A total of fourteen learners of Japanese participated in Kashiwagi (2011). However, to demonstrate a clear effect of working memory capacity, the results of only 10 participants with 4 participants removed from the median of L2 Japanese reading span test scores were analyzed.

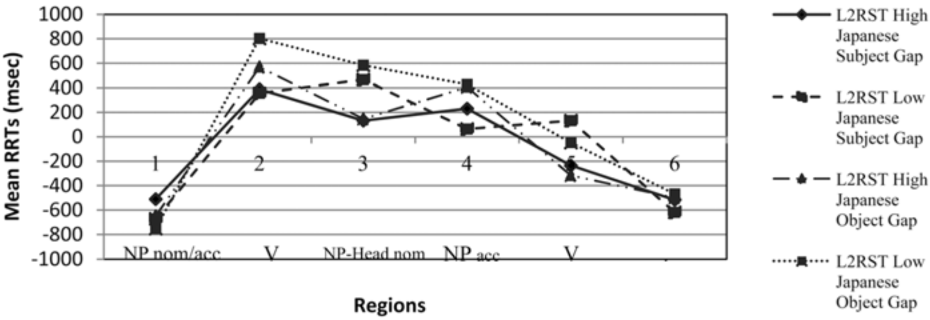



**Figure 5:** Mean residual reading times of Japanese gap sentences by RST groups of ten L2 learners (Kashiwagi 2011: 101)

was found in Region 3 (*ozisan ga*) and in the Japanese object-gap sentences, the difference was found in Region 5 (*hirotta*) by item. Findings also revealed that the L2 RST High and Low groups process these two types of sentences differently. The analyses of the L2 RST High group showed a significant difference between the Japanese subject-gap and object-gap sentences in Region 5 (*hirotta*) and differences for the L2 RST Low group in Region 2 (*miokutta*) and Region 4 (*omotya o*) by item. Therefore, Kashiwagi's results suggest the influence of working memory capacity on the reading times of Japanese relative clause sentences by L2 learners of Japanese.

Kashiwagi (2011) also compared the sentence reading performance of 10 Japanese native speakers and 10 L2 learners of Japanese. Japanese native participants were also divided into two working memory capacity groups using the Japanese reading span test: L1 RST High and L1 RST Low groups. Figures 6 and 7 show distinctive reading patterns among the different L1 and L2 RST groups.

The findings show a significant difference in Region 3 among L1 RST High group vs. L1 RST Low, L2 RST High and L2 RST Low groups in the subject-gap sentences. However, no difference was found in the object-gap sentences.[18] Overall, Kashiwagi's study supports the effect of working memory capacity on L2 Japanese sentence processing. Additionally, her results indicate that L2 learners may process a certain type of sentence structure similarly to native speakers with lower working memory capacity. This result is similar to the result of Havik, Roberts, van Hout, Schreuder, and Haverkort (2009) discussed earlier. It indicates that L2 learners with higher working memory capacity may process sentences similarly to the native speakers with lower working memory capacity.

---

**18** Although the figure of object-gap sentences seems to show a large difference among the four RST groups, especially in Region 2, no significance was found in post-hoc Tukey's HSD comparisons.

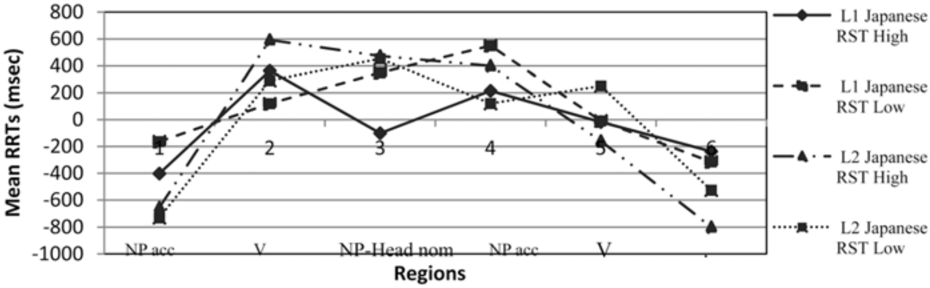



**Figure 6:** Mean residual reading times of Japanese subject-gap sentences by RST groups of ten L1 native speakers and ten L2 learners (Kashiwagi 2011: 124)

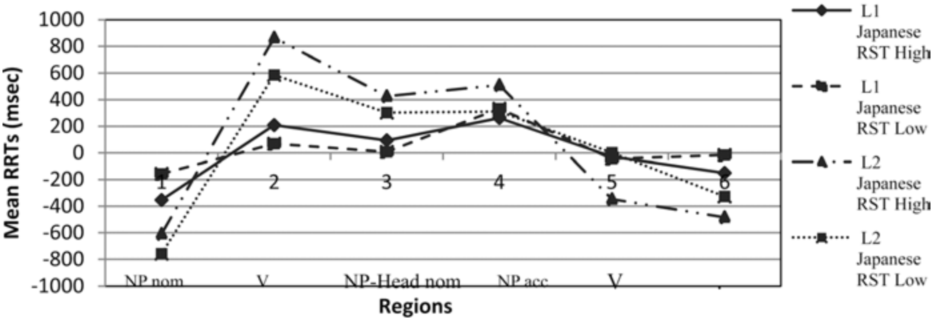



**Figure 7:** Mean residual reading times of Japanese object-gap sentences by RST groups of ten L1 native speakers and ten L2 learners (Kashiwagi 2011: 126)

This section has discussed the role of working memory in sentence processing in L2 Japanese. It is indicated that working memory may play a crucial role in L2 processing and the use of L2. However, more research needs to be done in order to fully understand the role and function of working memory as well as its constraints on L2 sentence processing. For example, in this section, we mainly dealt with L2 online sentence processing, but some studies suggest that the working memory effect may be much greater in other aspects of L2 such as discourse comprehension and vocabulary learning. Since L2 learners' proficiency develops over time, it would also be valuable to examine how the effect of working memory changes over the course of L2 development. Moreover, it is important to develop universal measurements that can be used for L2 Japanese learners with different backgrounds and to evaluate their reliability and validity. By finding one piece of evidence at a time, studies on working memory and L2 sentence processing can make important contributions to the field of L2 studies.

# 4 Conclusion and future research

In this chapter, we focused our discussion on L2 Japanese sentence processing and reviewed the relevant literature. The fundamental question, whether L2 sentence processing is different from L1 sentence processing, was discussed. By examining specific sentence structures, namely relative clause constructions and simplex sentence constructions with no garden-path or ambiguity, and incrementality in L2 sentence processing, we showed that the answer to the fundamental question stated above is, yet, inconclusive. Some studies found that novice level learners tend to have nonnative-like processing, but as they become more proficient in the L2, some studies found that it is possible that L2 processing becomes more native-like. Another finding suggested that L2 processing cannot be completely identical with L1 processing even among proficient learners. The studies we reviewed also suggested that experimental sentences as well as the learners' L1 affects L2 sentence processing performance. Additionally, the possibility of the effect of working memory on L2 sentence processing, especially to those who have higher working memory capacity, leading to similar sentence processing manner as native speakers, was indicated. Although there is no concrete conclusion that can be made to answer the fundamental question, we can say with impunity that L2 sentence processing is an intricate process with factors such as learners' L1, proficiency level, working memory, etc. intertwining together and affecting the learners' processing.[19]

The field of L2 sentence processing is exciting because there are still many issues to be examined. Studies related to issues other than the topics discussed above include, for example, the role of prosodic knowledge on L2 sentence processing (e.g., Goss and Nakayama 2011; Shibata and Hurtig 2008). For example, Goss and Nakayama (2011) examined whether (and if so where) English-speaking L2 learners of Japanese insert a pause when reading syntactically ambiguous and unambiguous sentences (12a and 12b).

(12) a. (Unambiguously Right Branching)

| | | | |
|---|---|---|---|
| *ookii* | [*natu* | *no* | *miitingu*] |
| large | summer | GEN | meeting |

'the large meeting in the summer'

b. (Unambiguously Left Branching)

| | | | |
|---|---|---|---|
| [*yasui* | *apaato* | *no*] | *soto* |
| cheap | apartment | GEN | outside |

'outside of the cheap apartment'

19 Studies that take into account the theory of Competition Model explain the effects of L1 transfer in processing and production. However, we did not focus on those in this chapter because they have been extensively introduced elsewhere (e.g., Koda 2004; MacWhinney and Bates 1989; Sasaki 2003, volume 8(3) of *Applied Psycholinguistics* 1987).

They found that the learners were frequently unable to produce the correct prosody. Furthermore, they also found that L2 learners comprehended the meaning with the intended interpretation independently of their prosodic break. This finding indicates that L2 learners may not be using prosodic information for sentence comprehension similarly to native speakers and this issue requires further investigation. Other issues may include familiarity of the words and orthographies, the learners' L1, age, and overall linguistic exposure to the language. Moreover, recent research efforts in the field of psycholinguistics and neurolinguistics using brain-imaging techniques to examine brain activity responses, such as the functional magnetic resonance imaging (fMRI) and event related potentials (ERPs), have shed additional light on L2 sentence processing. However, this research has not investigated L2 Japanese learners since most studies are done with L2 English learners. For example, Guo, Guo, Yan, Jiang, and Peng (2009) examined L2 English learners and verb sub-categorization violations, finding that different strategies were used by native speakers and L2 learners. The fMRI allows researchers to see the neural activity in participants' brains and tracks imaging changes depending on the blood flow related to energy used in the brain cells. Using this technique, for example, Buchweitz, Mason, Hasegawa and Just (2009) compared the brain activations of Japanese native speakers in their L1 (Japanese) and their L2 (English). They found that L2 reading requires more effort than L1 and that different brain responses were obtained between Japanese orthographies and the English alphabet. Another study by Yokoyama Okamoto, Miyamoto, Yoshimoto, Kim, Iwata, Jeong, Uchida, Ikuta, Sassa, Nakamura, Horie, Sato and Kawashima (2006) investigated whether L1 and L2 learners' process structurally complex sentences in Japanese and English differently. Using these new methods in L2 Japanese studies could be promising because they can provide evidence from different angles as well as cross-linguistic evidence. Moreover, examining a wider range of learners will yield a more complete picture of the factors that affect L2 sentence processing. Since only a limited number of studies have been done on L2 Japanese sentence processing, there is a great potential for future research and valuable contributions to the L2 field as well as to our overall understanding of human cognitive architecture.

## Acknowledgments

We thank Mineharu Nakayama for his patient, continuous comments and encouragements to the earlier versions of our drafts. We also thank anonymous reviewers for their helpful comments. All the shortcomings of the final draft are of course ours.

# References


Ahrens, Kathleen. 2003. Verbal integration: The interaction of participant roles and sentential argument structures. *Journal of Psycholinguistic Research* 32(5). 497–516.

Alptekin, Cem and Gülcan Erçetin. 2010. The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading* 33(2). 206–219.

Altmann, Gerry T. M. and Yuki Kamide. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73. 247–264.

Baddeley, Alan D. 2003. Working memory and language: An overview. *Journal of Communication Disorders* 36. 189–208.

Baddeley, Alan D. and Graham Hitch. 1974. Working memory. In Gordon H. Bower (ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8), 47–89. New York: Academic Press.

Baddeley, Alan D. and Robert H. Logie. 1999. Working memory: The multiple component model. In Akira Miyake and Priti Shah (eds.), *Models of working memory: Mechanisms of active maintenance and executive control*, 28–61. New York: Cambridge University Press.

Boland, Julie E. and Allison Blodgett. 2006. Argument status and PP-attachment. *Journal of Psycholinguistic Research* 35. 385–403.

Buchweitz, Augusto, Robert A. Mason, Mihoko Hasegawa and Marcel A. Just. 2009. Japanese and English sentence reading comprehension and writing systems: An fMRI study of first and second language effects on brain activation. *Bilingualism: Language and Cognition* 12. 141–151.

Caplan, David and Gloria S. Waters. 1995. On the nature of the phonological output planning process involved in verbal rehearsal: Evidence from aphasia. *Brain and Language* 48. 191–220.

Caplan, David and Gloria S. Waters. 1999. Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences* 22. 77–126.

Caplan, David and Gloria S. Waters. 2002. Working memory and connectionist models of parsing: A reply to MacDonald and Christiansen (2002). *Psychological Review* 109(1). 66–74.

Chipere, Ngoni. 2003. *Understanding complex sentences: Native speaker variation in syntactic competence*. New York, NY: Palgrave Macmillian.

Chun, Dorothy M. and Scott J. Payne. 2004. What makes students click: Working memory and look-up behavior. *System* 32(4). 481–503.

Clahsen, Herald. 2011. Grammatical processing in late bilinguals. Paper presented at Generative Approach To Language Acquisition (GALA), Capsis Hotel in Thessaloniki, Greece, 6–8 September.

Clahsen, Herald and Claudia Felser. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27(1). 3–42.

Conway, Andrew R.A., Michael J. Kane, Michael F. Bunting, Donald Z. Hambrick, Oliver Wilhelm and Randall W. Engle. 2005. Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review* 12. 769–786.

Currah, Satomi. 2004. *Processing Japanese adnominal structures: An empirical study of native and non-native speakers' strategies*. Edmonton, Alberta: University of Alberta dissertation.

Daneman, Meredyth and Patricia A. Carpenter. 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19. 450–466.

Daneman, Meredyth and Ian Green. 1986. Individual differences in comprehending and producing words in context. *Journal of Memory and Language* 25. 1–18.

Daneman, Meredyth and Phil M. Merikle. 1996. Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review* 3. 422–433.

Dussias, Paola and Pilar Piñar. 2009. Sentence parsing in L2 learners: Linguistic and experienced-based factors. In William C. Ritchie and Tej K. Bhatia (eds.), *The new handbook of second language acquisition*, 296–318. Bingley: Emerald Group Publishing.

Fedorenko, Evelina, Edward Gibson and Douglas Rohde. 2004. Verbal working memory in sentence comprehension. Paper presented at 26th Annual Meeting of Cognitive Science Society, Westin River North Hotel in Chicago, 5–7 August.

Felser, Claudia and Leah Roberts. 2007. Processing *wh*-dependencies in a second language: A cross-modal priming study. *Second Language Research* 23(1). 9–36.

Fortkamp, Mailce B.M. 1999. Working memory capacity and aspects of L2 speech production. *Communication and Cognition* 32. 259–296.

Frazier, Lyn and Charles Clifton Jr. 1996. *Construal*. Cambridge, MA: MIT Press.

Frazier, Lyn and Keith Rayner. 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* 14. 178–210.

Friedman, Naomi P. and Akira Miyake. 2004. The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language* 51. 136–158.

Fukuda, Michiko. 2004. Listening comprehension in Japanese as a second language and the capacity of working memory by adult-Malay-language-speaking learners. *Acquisition of Japanese as a Second Language* 7. 45–59.

Fukuda, Michiko. 2005. Listening comprehension in Japanese as a second language and working memory span: A study on differences in proficiency, targeting native speakers of Chinese. *Bulletin of the Graduate School of Education, Hiroshima University, Arts and Science Education* 2(53). 299–309.

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Alec P. Marants, Yasushi Miyashita and Wayne O'Neil (eds.), *Image, language brain: Papers from the first Mind Articulation Project Symposium*, 95–126. Cambridge, MA: MIT Press.

Goss, Seth and Mineharu Nakayama. 2011. Prosody and comprehension in oral reading by L2 Japanese learners. *Second Language* 10. 33–50.

Guo, Jingjing, Taomei Guo, Yan Yan, Nan Jiang and Danling Peng. 2009. ERP evidence for different strategies employed by native speakers and L2 learners in sentence processing. *Journal of Neurolinguistics* 22. 123–134.

Harrington, Michael W. and Mark Sawyer. 1992. L2 working memory capacity and L2 reading skills. *Studies in Second Language Acquisition* 14(1). 25–38.

Havik, Else, Leah Roberts, Roeland van Hout, Robert Schreuder and Marco Haverkort. 2009. Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning* 59(1). 73–112.

Ikeno, Osamu. 2006. L1 and L2 working memory: An investigation into the domain specificity and processing efficiency issues. *Bulletin of the Faculty of Education Ehime University* 53(1). 113–121.

Itomitsu, Masayuki and Mineharu Nakayama. 2005. Reading span test (English L1) and RST as a measuring tool for working memory. Unpublished ms., The Ohio State University, Columbus, OH.

The Japan Foundation. 2011. *Survey report on Japanese-language education abroad 2009*. Tokyo: The Japan Foundation.

Jincho, Nobuyuki and Reiko Mazuka. 2011. Individual differences in sentence processing: Effects of verbal working memory and cumulative linguistic knowledge. In Hiroko Yamashita, Yuki Hirose, and Jerome L. Packard (eds.), *Processing and producing head-final structures, Studies in Theoretical Psycholinguistics*, 49–65. Dordrecht, Heidelberg, London, and New York: Springer.

Juffs, Alan. 2000. An overview of the second language acquisition of the links between verb semantics and morpho-syntax. In John Archibald (ed.), *Second Language Acquisition and Linguistic Theory*, 187–227. Oxford: Blackwell.

Juffs, Alan. 2002. Working memory as a variable in accounting for individual differences in second language performance. American Association of Applied Linguistics Annual Meeting, Salt Lake City, 6–9 April.

Juffs, Alan. 2004. Representation, processing and working memory in a second language. *Transactions of the Philological Society* 102. 199–226.
Juffs, Alan. 2005. The influence of first language on the processing of *wh*-movement in English as a second language. *Second Language Research* 21. 121–151.
Juffs, Alan and Michael W. Harrington. 2011. Aspects of working memory in L2 Learning. *Language Teaching: Reviews and Studies* 42(2). 137–166.
Just, Marcel Adam and Patricia A. Carpenter. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review* 99(1). 122–149.
Kahraman, Bariş. 2011. *Nihongo oyobi Torukogo ni okeru "kusho to maigo no izon kankei" no shori: Bunshori no chikujisei o megutte* [Processing "gap-filler dependencies" in Japanese and Turkish: Regarding the incrementality of sentence processing]. Hirohima, Japan: Hiroshima University dissertation.
Kahraman, Bariş. 2012. Sentence processing of nominative-genitive conversion in Japanese by Turkish speaking learners and native speakers. Paper presented at the 13th Tokyo Conference on Psycholinguistics, Keio University, Tokyo, 9–10 March.
Kahraman, Bariş and Hiromu Sakai. 2015. Relative clause processing in Japanese: Psycholinguistic investigation into typological differences. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.
Kahraman, Bariş, Atsushi Sato, Mariko Koide, Mariko Uno, Miwa Takemura and Hiromu Sakai. 2009. Processing Japanese subject and object relative clauses by advanced learners: Comparison with native speakers by a whole-sentence reading experiment. *IEICE Technical Report* 18. 57–62.
Kamide, Yuki, Gerry T. M. Altmann and Sarah L. Haywood. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye-movements. *Journal of Memory and Language* 49. 133–156.
Kanno, Kazue. 2001. On-line processing of Japanese by English L2 learners. *Acquisition of Japanese as a Second Language* 4. 23–38.
Kanno, Kazue. 2007. Factors affecting the processing of Japanese relative clauses by L2 learners. *Studies in Second Language Acquisition* 29. 197–218.
Kashiwagi, Akiko. 2011. *Processing relative clauses in first and second language*. Columbus, OH: The Ohio Sate University dissertation.
Keating, Gregory. 2009. Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning* 59. 503–535.
Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8. 63–99.
King, Jonathan and Marcel Adam Just. 1991. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language* 30. 580–602.
Koda, Keiko. 2004. *Insights into second language reading*. Cambridge: Cambridge University Press.
Kwon, Nayoung, Maria Polinsky, and Robert Kluender. 2006. Subject preference in Korean. In Donald Baumer, David Montero and Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conferences on Formal Linguistics*, 1–14. Somerville, MA: Cascadilla Proceedings Project.
Lieberman, Moti, Sachiko Aoshima, and Colin Phillips. 2006. Native-like biases in generation of *wh*-questions by non-native speakers of Japanese. *Studies in Second Language Acquisition* 28. 423–448.
Lin, Chen-Jer Charles and Thomas G. Bever. 2006. Subject preference in the processing of relative clauses in Chinese. In Donald Baumer, David Montero and Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conferences on Formal Linguistics*, 254–260. Somerville, MA: Cascadilla Proceedings Project.
MacDonald, Maryellen C. and Morten H. Christiansen. 2002. Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review* 1. 35–54.

MacDonald, Maryellen C., Marcel A. Just and Patricia A. Carpenter. 1992. Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology* 24. 56–98.
MacWhinney, Brian and Elizabeth Bates (eds.). 1989. *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press.
Mazuka, Reiko, Kenji Itoh and Tadahisa Kondo. 2002. Costs of scrambling in Japanese sentence processing. In Mineharu Nakayama (ed.), *Sentence processing in East Asian languages*, 131–166. Stanford, CA: CSLI Publications.
McElree, Brian. 2006. Accessing recent events. *The Psychology of Learning and Motivation* 46. 155–200.
Mitsugi, Sanako. 2011. *The use of morphosyntactic and lexical semantic information during sentence processing in second language Japanese*. Pittsburgh, PA: Carnegie Melon University dissertation.
Mitsugi, Sanako, Brian MacWhinney and Yasuihiro Shirai. 2010. Cue-based processing of relative clauses in L2 Japanese. In Matthew T. Prior, Yukiko Watanabe, and Sang-Ki Lee (eds.), *Selected Proceedings of the 2008 Second Language Forum*, 123–138. Somerville, MA: Cascadilla Proceedings Project.
Miyake, Akira and Naomi P. Friedman. 1998. Individual differences in second language proficiency: Working memory as language aptitude. In Alice F. Healy, and Lyle E. Bourne, Jr. (eds.), *Foreign language learning: Psycholinguistic studies on training and retention*, 339–364. Mahwah, NJ: Lawrence Erlbaum.
Miyake, Akira, Marcel A. Just and Patricia A. Carpenter. 1994. Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts. *Journal of Memory and Language* 33. 175–202.
Miyake, Akira and Priti Shah (eds.). 1999. *Models of working memory: Mechanisms of active maintenance and executive control.* New York: Cambridge University Press.
Miyamoto, Edson T. 2002. Case marker as clause boundary inducers in Japanese. *Journal of Psycholinguistic Research* 31(4). 307–347.
Miyamoto, Edson T. and Michiko Nakamura. 2003. Subject/object asymmetries in the processing of relative clauses in Japanese. In Gina Garding and Mimu Tsujimura (eds.), *Proceedings of the 22nd West Coast Conferences on Formal Linguistics*, 342–355. Somerville, MA: Cascadilla Press.
Nakano, Yoko, Claudia Felser and Herald Clahsen. 2002. Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research* 31. 531–571.
O'Grady, William. 1999. Toward a new nativism. *Studies in Second Language Acquisition* 21. 621–633.
O'Grady, William. 2001. A linguistic approach to the study of language acquisition. *Journal of the Pan-Pacific Association of Applied Linguistics* 5. 57–71.
Omaki, Akira. 2005. Working memory and relative clause attachment in first and second language processing. Honolulu, HI: University of Hawai'i master's thesis.
Osaka, Mariko. 2000. Waakingu memori to gengo rikai no nounai kouzou [Brain architecture for working memory and language comprehension]. In Naoyuki Osaka (ed.), *Noo to waakingu memori* [Brain and working memory], 157–180. Kyoto: Kyoto Daigaku Gakujutsu Shuppankai.
Osaka, Mariko. 2002. *Noo no memochou: Waakingu memori* [Notepad in brain: Working memory]. Tokyo: Shinyousha.
Osaka, Mariko and Naoyuki Osaka. 1992. Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society* 30(4). 287–289.
Osaka, Mariko and Naoyuki Osaka. 1994. Working memory capacity related to reading: Measurement with the Japanese version of reading span test. *The Japanese Journal of Psychology* 65 (5). 339–345.

Osaka, Mariko and Naoyuki Osaka. 2002. Individual differences in working memory during reading with and without parafoveal information: A moving-window study. *The American Journal of Psychology* 115(4). 501–513.

Osaka, Mariko, Naoyuki Osaka and Rudolf Groner. 1993. Language-independent working memory: Evidence from German and French reading span tests. *Bulletin of the Psychonomic Society* 31(2). 117–118.

Ozeki, Hiromi. 2008. *The acquisition process of Japanese noun-modifying clauses by first and second language learners*. Tokyo: Kurosio Publishers.

Ren, Marvin Hulin. 2009. Working memory and Chinese learners' processing of complex English sentences. *Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching* 3. Retrieved from http://www.ling.lancs.ac.uk/pgconference/v03.htm (accessed 17 November 2012)

Roberts, Michael Alexander. 2000. *Implicational markedness and the acquisition of relativization by adult learners of Japanese as a foreign language*. Honolulu, HI: University of Hawai'i dissertation.

Roberts, Rose and Edward Gibson. 2002. Individual differences in sentence memory. *Journal of Psycholinguistic Research* 31. 573–598.

Rodriguez, Guillermo A. 2008. *Second Language Sentence Processing. Is it fundamentally different?* Pittsburgh, PA: University of Pittsburgh dissertation.

Sagarra, Nuria and Julia Herschensohn. 2010. The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua* 120(8). 2022–2039.

Sasaki, Yoshinori. 2003. Contentions of second language acquisition research based on the competition model: A perspective from JSL. In Yukiko A. Hatasa (ed.), *An invitation to second language acquisition research in Japanese: In honor of Seiichi Makino*, 155–169. Tokyo: Kurosio Publishers.

Sato, Atsushi. 2011. *Nihongo kankeisetsu no shorifuka o kettei suru yoin no kento: Koopasu ni okeru shiyohindo no eikyo o chushin ni* [A study on determining factors for processing load of Japanese relative clauses: Influence of corpus-based frequency]. Hiroshima, Japan: Hiroshima University dissertation.

Sawasaki, Koichi. 2004. L2 sentence processing by English-speaking learners of Japanese. *Ars Linguistica* 11. 136–155.

Sawasaki, Koichi. 2007. *L2 reading by learners of Japanese: A comparison of different L1s*. Columbus, OH: The Ohio State University dissertation.

Sawasaki, Koichi. 2008. Eigo o bogo to suru L2 nihongo gakushusha no kankeisetsu rikai: Yomijikan kara no kosatsu [Comprehension of relative clauses by English-speaking L2 learners of Japanese]. *Ars Linguistica* 15. 1–20.

Sawasaki, Koichi. 2009a. Processing of relative clauses by learners of Japanese: A study on reading times of English/Korean/Chinese L1 speakers. *Acquisition of Japanese as a Second Language* 12. 86–106.

Sawasaki, Koichi. 2009b. Meishiko renzoku to ga-kaku renzokubun no na'nido to memorii supan no kankei [Working memory span and its effect on difficulty ratings of sentences with repeated NP arguments and repeated NP-ga]. *Ars Linguistica* 16. 15–31.

Sawasaki, Koichi. 2012. Nihongo bunshori ni tsuite: L1 to L2 no kankeisetsu shori o chushin ni [On Japanese sentence processing: Processing of L1 and L2 relative clauses]. A talk delivered at J-SLA Autumn Seminar 2012 (The Japan Second Language Association), Chuo University, Tokyo, 28 October.

Shibata, Tomoko and Richard R. Hurtig. 2008. Prosody acquisition by Japanese learners. In Zhao-Hong Hang (ed.), *Understanding second language process*, 176–204. Clevedon, UK: Multilingual Matters Ltd.

Tamaoka, Katsuo. 2015. Processing of the Japanese language by native Chinese speakers. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Tokimoto, Shingo. 2004. Reanalysis costs in processing Japanese sentence with complex NP structures and homonyms: Individual differences and verbal working memory constraints. *Japanese Cognitive Science Society: Technical Report* 53. 1–14. Retrieved from http://www.jcss.gr.jp/ (accessed 17 November 2012)

Trueswell, John C., Michael K. Tanehaus and Susan M. Garnsey. 1994. Semantic influence on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33(3). 285–313.

Trueswell, John C., Michael K. Tanehaus and Christopher Kello. 1993. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-path. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19(3). 528–553.

Ueno, Mieko and Susan M. Garnsey. 2007. Gap-filling vs. filling gaps: Event-related brain indices of subject and object relative clauses in Japanese. In Naomi Hanaoka McGloin and Junko Mori (eds.), *Proceedings of the 15th Japanese-Korean Linguistics Conference*, 286–298. Stanford, CA: CSLI Publications.

Van den Noort, Maurits, Bosch Peggy, Marco Haverkort and Kenneth Hugdahl. 2006. Foreign language proficiency and working memory capacity. *European Psychologist* 11(4). 289–296.

Walter, Catherine. 2004. Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics* 25(3). 315–339.

Wanner, Eric and Michael Maratos. 1978. An ATN approach to comprehension. In Morris Halle, Joan Bresnan and George A. Miller (eds.), *Linguistics theory and psychological reality*, 119–161. Cambridge, MA: MIT Press.

Watanabe, Fuyumi. 2012. Reading span test for Japanese language learners: Measuring working memory capacity in L2 reading. *Keiōgijjuku Daigaku Nihongo Nihonbunka Kyōiku Sentā Kiyō Nihongo to Nihongo Kyōiku* [Keio University Center for Japanese Studies Japanese and Japanese language education] 40. 113–119.

Waters, Gloria S. and David Caplan. 1996. The capacity theory of sentence comprehension critique of Just and Carpenter. *Psychological Review* 103. 1–12.

Yamashita, Hiroko. 1994. *Processing of Japanese and Korean*. Columbus, OH: The Ohio State University dissertation.

Yamashita, Hiroko. 1997. The effect of word-order and case marking information on the processing of Japanese. *Journal of Psycholinguistic Research* 26(2). 163–188.

Yamashita, Hiroko. 2008. Effects of sentence processing strategy proximity on the comprehension of second languages. *Second Language* 7. 43–82.

Yokoyama, Satoru, Hideyuki Okamoto, Tadao Miyamoto, Kei Yoshimoto, Jungho Kim, Kazuki Iwata, Hyeonjeong Jeong, Shinya Uchida, Naho Ikuta, Yuko Sassa, Wataru Nakamura, Kaoru Horie, Shigeru Sato and Ryuta Kawashima. 2006. Cortical activation in the processing of passive sentences in L1 and L2: An fMRI study. *Neuroimage* 30. 570–579.

Zhai, Yong. 2009. Developmental shift of parsing strategies: Processing empty subject sentences among L1 and L2 learners. Fukuoka, Japan: Kyushu University dissertation.

Noriko Iwasaki

# 17 Sentence production models to consider for L2 Japanese sentence production research

## 1 Introduction

Sentence production (i.e., cognitive processes engaged when one is speaking) is a "relatively young field of investigation" (Costa, Alario, and Sebastián-Gallés 2009: 531), compared to sentence comprehension/processing, due largely to methodological difficulties. While it is relatively easy to manipulate properties of stimuli (such as their complexity) for readers or listeners to comprehend and to then measure their behavior (e.g., response time), the manipulation of input for speakers, namely concepts or messages to convey, is not as straightforward. Since the 1990s, however, there has been a dramatic increase in research on representations and processes involved in language production, utilizing various experimental approaches from cognitive psychology (Bock 1996).

Research on cognitive processes involved in second language (L2) sentence production (i.e., speaking) is even younger. When it comes to research on L2 Japanese sentence production, the situation is even more taxing, as research on first language (L1) sentence production has been conducted in a very limited number of languages mostly belonging to the Germanic family (English, Dutch, and German) or to the Romance family (Spanish, Catalan, Italian and French) (Costa et al. 2009). In other words, we are not even certain to what extent the L1 sentence production models proposed so far reflect general language processes shared across typologically different languages, and hence we cannot assume that L2 sentence production models based on L1 models are general across different L2s, either. There may indeed be mechanisms that are specific to Japanese and/or languages that share some linguistic features with Japanese.

The number of published studies on L1 Japanese sentence production is a far cry from those on Germanic or Romance languages, as suggested by Yamashita and Chang's (2006) statement that the study of sentence production in Japanese "is still in infancy". Yet, Japanese is probably the most studied non-European language. In fact, Jaeger and Norcliffe (2009: 12) consider Japanese to be one of the languages for which there is a "sizable" literature on sentence production of "more than five papers". Studies utilizing the features of Japanese that diverge from or are absent in the languages primarily studied, such as verb-final word order, the relative flexibility of word order, and the case-marking system, recently shed light on aspects of sentence production, notably incremental processes to be briefly reviewed in this chapter.

Studies on Japanese have also contributed to other research areas related to language production research: language-specific and universal ways of describing motion events, which occur with co-speech gesture. This involves "tightly coordinated interactions between systems in different modalities" which is "ripe for vigorous investigation" (Schiller, Ferreira, and Alario 2007: 1147).

Another Japanese feature that might be of interest to researchers is the large number of mimetic, or sound-symbolic, words, which are quintessentially iconic (i.e., form-meaning mapping is motivated by resemblance between the word form and its referent). There is now ample evidence that "speakers and signers exploit iconicity in language processing" (Perniss, Thompson and Vigliocco 2010). In Japanese, mimetic adverbs co-occur with gesture when speakers describe events (Kita 1997; Kita 2001). Example (1) below is from Kita (1997: 393). The non-italicized mimetic word *baa* was accompanied by a stroke (a meaningful phase of a gesture), i.e., the right hand index finger is extended and moves forcefully downward) and the underlined *to* was accompanied by holding an arm in the air. The onset and the end of a gesture are indicated by square brackets.

| | | | | | |
|---|---|---|---|---|---|
| (1) | [*biru* | *o* | *baa* | <u>*to*</u> ] | *sagat-te* |
| | building | ACC | MIM | COMP | go.down-GER[1] |
| | '(the cat) goes down the building with great momentum, and' | | | | |

Though research on L2 Japanese language processes is very limited, developments in the areas above have implications for L2 Japanese production research. Hence, this chapter reviews sentence production models and empirical studies that are pertinent to L2 Japanese sentence production and future investigation. In Section 2, focusing on grammatical encoding (which includes lexical access), the "consensus" model of sentence production is first described, followed by summaries of studies specific to the Japanese language in order to reveal (potential) differences between sentence production processes in European languages and Japanese. In Section 3, L2 (bilingual) sentence production models (de Bot 1992; Hartsuiker, Pickering and Veltkamp 2004) based primarily on L1 and L2 research examining European languages are reviewed, and implications for L2 Japanese are discussed. In Section 4, studies on language-specific description of motion events with co-speech gesture are summarized, and a gesture-speech model is presented. In Section 5, recent findings on motion description among bilingual speakers who speak Japanese (or Korean, a language typologically similar to Japanese) as one of their languages are summarized. Finally concluding remarks are provided in Section 6.

**1** The notations used in glosses in this chapter are: ACC accusative, GEN genitive, GER gerund, LOC locative, NMLZ nominalizer, NOM nominative, COMP complementizer, MIM mimetic word, PASS passive.

# 2 Sentence production model and the Japanese language processes

## 2.1 Overview of the "consensus" sentence production model

While there are debates with regard to some aspects of the processes involved in sentence production (discussed in Section 2.2), there is largely consensus in the stages of processes (depicted in Figure 1), which are based on the seminal works by Garrett (1975, 1988) and Levelt (1989). The focus in these works is on grammatical encoding, "[t]he heart of the faculty of language" (Ferreira and Slevc 2007), including lexical access. (Conceptualization in relation to expressing motion events will also be discussed in Section 4.)

Grammatical encoding refers to processes in which preverbal messages (i.e., conceptual structures to be expressed) are developed into well-formed sentences. These processes proceed by first mapping the participants in the conceptualized events to their thematic roles, such as agent and patient, and assigning them grammatical functions (e.g., subject, object) at the functional level. Next, syntactic

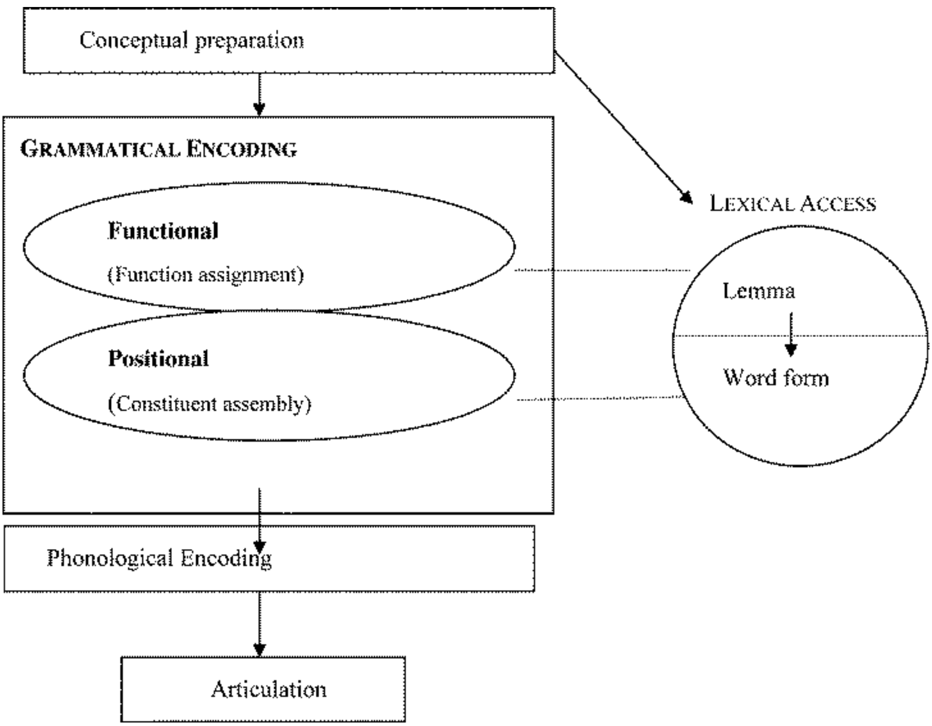



**Figure 1:** Levelt's and Garrett's models integrated

frames are built and constituents are ordered at the positional level. In other words, the functional-level processes map the speaker's thinking to language, and the positional-level processes map the linguistic representation prepared at the functional level to linear order. At the same time, lexical information is retrieved and integrated. The preverbal message contains the information necessary to retrieve the words via lexical concepts, i.e., concepts for which verbal labels are available (Levelt, Roelofs and Meyer 1999) or lexico-semantic representations that bind the distributed conceptual features (Vigliocco et al. 2004). Each lexical concept or lexico-semantic representation triggers access to lemma, which contains a package of syntactic information of the respective lexical entry, e.g., grammatical classes and subcategorization features, information indispensible to carrying out the function assignment. The selected lemma then activates its target word form (i.e., the morphological, segmental and metrical spell-outs of the word), which is then integrated/inserted into the frames built at the positional level.

## 2.2 Debates: The nature and directionality of information flow

The model above is characterized by the simultaneous working of two processes, structure building and lexical access, both of which further consist of two stages – functional and positional, and lemma and word form. The major debates among sentence production researchers concern the nature and directionality of information flow between stages as well as the interface between structure building and lexical access. Information flow is either serial/discrete or cascading, while directionality is either unidirectional or bidirectional, with the latter allowing for feedback. The original models proposed by Garrett (1975) and followed up by Levelt (1989) assume the stages to be unidirectional, serial and discrete, i.e., the process of a given minimal unit at one stage needs to be completed for its output to be sent to the next stage, and only minimal, necessary information is passed on. Levelt et al. (1999) also proposed a comprehensive model of lexical access that abides by these principles.

However, there is now good evidence that information flow is not strictly serial, and rather, more information is passed on to the next level without waiting to complete the process. Also, there is feedback from the later stage to the previous stage. Such evidence has been robust for word form retrieval (see Meyer and Belke 2007), but there is also substantial evidence for structure building as well (see Vigliocco and Hartsuiker 2002). We will see examples of a structure-building phenomenon, namely subject-verb agreement, which provides evidence for non-discrete information flow and feedback in this sub-section and revisit this issue with regard to Japanese in Sections 2.3 and 2.4.

Subject-verb number (and gender) agreement has been one of the major areas of investigation in sentence production research processes.[2] Experiments investigating errors of subject-verb agreement, functional-level processes, in European languages revealed that there is unnecessary information flowing from the conceptual-level to the functional-level, which nevertheless speakers utilize. NPs such as *the stamp on the envelopes* induced more subject-verb agreement errors (e.g., *The stamp on the envelopes ARE beautiful*) than NPs like *the trap for the rats* when experimental participants heard the NPs, repeated them and completed the sentences using the NPs as the subject of sentences (Eberhard 1999).[3] Despite the fact that both types of NPs have the singular noun as the head, the referent of the former is conceptually plural (i.e., the same stamp used for multiple envelopes). This kind of effect of conceptual number on subject-verb agreement is called the "distributivity effect", and was observed across many languages such as Dutch, English, French, Italian and Spanish (Hartsuiker, Kolk and Huinck 1999; Vigliocco, Butterworth and Garrett 1996; Vigliocco, Butterworth and Semenza 1995). Similar effects were also found for gender agreement across many languages (e.g., Vigliocco and Franck 1999).

At the same interface between conceptual preparation and functional-level, Vigliocco and Hartsuiker also suggest the possibility of feedback from functional level processes to conceptual preparation on the basis of Slobin's (1996) thinking-for-speaking hypothesis (discussed further in Section 4). According to this hypothesis, conceptual preparation differs depending on what information needs to be grammatically encoded in the language spoken. For example, English speakers need to prepare conceptual structures that specify number (singular or plural) while Japanese speakers do not. Thus Vigliocco and Hartsuiker (2002: 457) argue that "feedback from phrasal processes may fine-tune the conceptual representation to speaking" though Levelt (1989) considered this to be applicable only to language learners.

Previous studies on Japanese further suggest information flow that was not found in studies on other languages. In the next section, we first summarize studies on aspects of structure building (i.e., the effect of conceptual accessibility on structure building) and speech errors of case particles. This is followed by studies related to lexical access. Note that lexical access, especially the selection of lemmas, is crucial for structure building as the model depicted in Figure 1 suggests, and thus is an important sub-process of grammatical encoding.

---

**2** Subject-verb agreement is considered to be a significant phenomenon in research on sentence production since it allows researchers to investigate "[a] central question in language production" that relates to how intrinsic syntactic dependencies between words in sentences are computed (Costa et al. 2007: 536). However, since languages like Japanese do not have subject-verb agreement, reliance on research on this phenomena is indicative of why models proposed so far may not necessarily account for sentence production processes in Japanese.

**3** Earlier studies on subject-verb agreement in English (Costa 2005) did not show this effect, but Eberhard's (1999) study showed that the absence of the effects in earlier studies might have been due to the types of stimuli used in the experiments, whose referents were more difficult to imagine than the stimuli Eberhard used.

## 2.3 Conceptual accessibility and structure building

Experiments on Japanese provided evidence that conceptual information affects structure building – both grammatical role assignment (the functional level) and linear order (the positional level). Some conceptual properties make some entities more accessible than others. Such properties include salience/prominence, animacy, the status of information as given or new, and imageablity (i.e. probability of evoking strong image). McDonald, Bock and Kelly (1993) found that English speakers prefer to have animate NPs as the subject by choosing passive or active transitive sentences (e.g. *A farmer purchased a refrigerator; The child was soothed by the music*). This was interpreted as the effect of conceptual accessibility on grammatical role assignment (functional level processes) since animacy was not found to affect the ordering of two nouns in conjunction (e.g., *the camera and crew*). Because the word order of English is rather rigid, it is not possible to examine the effect of animacy on the word order of agent NPs and patient NPs independent of grammatical roles (subject/object).

Utilizing the flexibility of word order in Japanese, Tanaka et al. (2011) investigated the effect of animacy on Japanese speakers' choice of passive/active and canonical/scrambled sentences in sentence recall tasks. Not only did they replicate the animacy effect on passive/active voice in Japanese, but they also found the effect on linear word order. Japanese speakers preferred to have the animate NPs as first NP in the sentence and as the grammatical subject. When Japanese speakers (N = 72) recalled sentences such as (2)–(5) that they heard earlier, they switched the word order of the OSV sentences like (3) to the canonical SOV order more often than for sentences like (2) (104 as compared to 46 instances out of 378 trials each)[4], and they also switched the voice of the passive sentences like (4) to active sentences, resulting in the animate NP *ryoosi* 'fisherman' as the subject NP, than for sentences like (5) (79 as compared to 38 instances out of 378 trials each).

(2)

| | | | | | | |
|---|---|---|---|---|---|---|
| *Minato* | *de* | *ryoosi* | *o* | *booto* | *ga* | *hakonda.* |
| harbor | LOC | fisherman | ACC | boat | NOM | carried |

'At the harbor, the boat carried the fisherman.'

(3)

| | | | | | | |
|---|---|---|---|---|---|---|
| *Minato* | *de* | *booto* | *o* | *ryoosi* | *ga* | *hakonda.* |
| harbor | LOC | boat | ACC | fisherman | NOM | carried |

'At the harbor, the fisherman carried the boat.'

(4)

| | | | | | | |
|---|---|---|---|---|---|---|
| *Minato* | *de* | *ryoosi* | *niyotte* | *booto* | *ga* | *hakobareta.* |
| harbor | LOC | fisherman | by | boat | NOM | carry.PASS |

'At the harbor, the boat was carried by the fisherman.'

4 Tanaka et al. (2011) report these figures in their Table 6. The total numbers of trials, 378, was calculated by the current author by summing the occurrences of all recall response types.

(5) *Minato de booto niyotte ryoosi ga hakobareta.*
harbor LOC boat by fisherman NOM carry.PASS
'At the harbor, the fisherman was carried by the boat.'

The mechanisms that cause these patterns to arise are not easy to explain, especially if we maintain the postulation of two stages (see Chang 2009 and this volume) because animacy is affecting both functional and positional processes. It is also not entirely clear whether the effect of animacy on word order in Japanese reflects processes that are fundamentally different from most European languages studied so far (except for Greek, studied by Branigan and Feleki 1999), or if the rigid word order makes it impossible to observe the effect of conceptual accessibility on word order in languages like English. But it is likely that the flexibility of word order promotes fluency, maximally allowing incremental processes (Kempen and Hoenkamp 1987; Levelt 1989). Speakers process easily accessible concepts as soon as they become available, assigning the grammatical role (the preferred grammatical role, being the subject-NP) to the lemma that is retrieved first. The word form that is retrieved first, which is likely to be the word form of the lemma that became available first, is then assigned the sentence-initial position.

## 2.4 What case particle errors tell us about structure building

The flexibility of word order comes with another feature of Japanese, namely, case marking by case particles. Speech errors (slips of the tongue) involving case particles also suggest that the flexibility of word order in Japanese allows speakers to process the language incrementally perhaps more so than speakers of rigid word order languages.

Case assignment is assumed to take place at the functional level and the assigned cases are realized at the positional level. Bock and Levelt (1994: 962) explain (6), a speech error of case in English (which has overt case only on pronouns) from Stemberger (1982), as an error due to mishap at the functional level (see also Melinger, Pechmann and Pappert 2009). The two pronouns, the second-person pronoun and the third-person-plural pronoun, are exchanged, and their lemmas are assigned unintended grammatical functions, but the case realized for each pronoun is correct for its respective position in the syntactic frame built at the positional level because they are intrinsic features of the frame.

(6) *You must be too tight for them.*
Intended: They must be too tight for you.

Japanese structural cases, such as nominative, accusative, and dative, are also assumed to be assigned at the functional level.[5] Each NP assigned a grammatical

---

**5** The distinction between structural case markers and postpositions is not very straightforward. See Sadakane and Koizumi (1995) about the syntactic status of *ni* and Inoue (1998) about *o*.

role may be tagged with the specification for case-marking realization, and the specific case marking of each case is then realized at the positional level. Meaningful case particles, or postpositions, such as instrumental *de* 'by', *kara* 'from' and *made* 'up to/until', which are comparable to English prepositions, are considered to be selected by their respective lexical concepts. Example (7) contains the nominative *ga* as well as three postpositions, *kara*, *made*, and *de*.

| | | | | | | |
|---|---|---|---|---|---|---|
| (7) | *Taroo* | *ga Tookyoo* | *kara Kyooto* | *made* | *Shinkansen de* | *itta.* |
| | | NOM | from | as.far.as | bullet.train by | went |

'Taro went from Tokyo to Kyoto by Bullet train.'

In German, speech errors of case marking are very rare. Berg (1987: 285) states that because case information is "not inherent to the moving nouns but assigned to it via the syntactic structure, it is not often involved in errors". However, case marking errors are not uncommon in natural Japanese spoken error corpora. For example, in his corpus consisting of 3200 speech errors, Terao (1995) found 373 case-marker and particle errors, of which 100 were errors of using *ga* where another marker should have been used such as (8) (Terao 1995: 253). Note that the erroneous elements are in capital letters below.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (8) | *Sassoku* | *hagaki* | *GA* | *yon-de* | *mi-tai* | *to* | *omoimasu.* |
| | right.away | postcard | NOM | read-GER | try-want | COMP | think |

'I would like to read a postcard right away'.

Iwasaki (2007) examined speech errors of case particles (both structural case markers and postpositions) experimentally elicited by a picture description task. After a brief presentation of each picture on the computer screen, one of the participants/entities was highlighted with color, and the participants were instructed to describe the picture starting their sentence with the highlighted participant/entity, mentioning all participants/entities in the event. When only argument NPs were considered, there was a total of 53 case particle errors, of which the nominative *ga* was the most frequently used erroneous particle: *ga* [18 errors] > *o* [11] > *no* [9] > *de* [3] > *ni* [2]. In (9), the patient NP *koohii* 'coffee' is marked with *ga* and repaired by the speaker.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (9) | *Koohii* | *GA* | *a,* | *koohii* | *o* | *weitaa* | *ga* | *kobosimasita.* |
| | coffee | NOM | (filler) | coffee | ACC | waiter | NOM | spilled(transitive) |

'The waiter spilled the coffee.'

The participants were also often found to adjust the rest of the sentence after articulating the *ga*-marked sentence-initial NP, which led to a grammatical sequence as seen in (10).

(10) *Koohii ga… koboreta node weitaa ga*
coffee NOM spilled(intransitive) because waiter NOM
*okyakusan ni ayamatta.*
customer to apologized
'Because the coffee spilled, the waiter apologized to the customer.'

The participants often resorted to constructing a subordinate clause such as (10), or passive construction such as (11), when participants started with the patient NPs marked by *ga*, suggesting their preference to mark the first retrieved noun form with *ga* and then generate or edit the rest.

(11) *Osara ga otokonoko niyotte warare…*
dish NOM boy by break.PASS
'The dish was broken by the boy (and…)'

Examining the same dataset, Iwasaki (2011) found that Japanese speakers also made more errors of *o* when the subject NP was the patient NP (i.e., the subject of passive or unaccusative verb predicates) than when it was the agent NP (i.e., the subject of active transitive or ergative sentences) as in (12)–(13). In (12), an unaccusative verb *tuiraku suru* 'fall' is used in the predicate, and in (13) the verb *dakko suru* 'hold' is passivized.

(12) *Hikooki O umi no naka ni tuiraku-sita.*
airplane ACC ocean GEN inside LOC fell
'The plane fell into the ocean.'

(13) *Onnanoko O okaasan ni dakko sare-te iru.*
girl ACC mother by hold PASS-GER is
'The girl is held by her mother.'

This finding indicates that speakers refer to the conceptual information (in this case, thematic role in the event) for case assignment. Some naturally occurring errors found by Iwasaki (2006a) also suggest that Japanese speakers refer to the nature of conceptualized events, namely, degree of transitivity (Hopper and Thompson 1980), when assigning case to NPs. For instance, NPs used to describe events with relatively high transitivity (e.g., tackling one's tasks) yet requiring *ni* for the object NP (*monogoto ni taioo suru* 'deal with/tackle') occurred with *o*, while NPs used to describe events with relatively low transitivity (e.g., supporting someone) requiring *o* for the object NP (*kare o ooen suru* 'support/cheer him') occurred with *ni.*

These findings suggest that case marking during Japanese sentence production does not always rely on the syntactic properties of the verb (e.g., subcategorization)

or predicates (e.g., passive/active). Instead, Japanese speakers appear to proceed by such mechanisms as assigning *ga* to the sentence-initial (argument) NPs or by referring to the conceptual information of the event (i.e., transitivity; the association between thematic role and case particle such as patient NP and the accusative *o*), to allow incremental processes maximally.

Perhaps in languages like Japanese, which do not have subject-verb agreement, the subject NP and the predicate are not as strongly linked, and a unit/scope of grammatical encoding can be smaller, which can also facilitate incremental processes. This possibility is further supported by findings that relate to the retrieval of verb lemmas as discussed below.

## 2.5 Lexical access (lemma and word form)

We are concerned here with lexical access during sentence production, that is, processes of selecting a word, accessing its syntactic information (lemma) and retrieving its word form. Both the original model that today's psycholinguists primarily drew from (Garrett 1975) and the models that followed it (Levelt 1989; Levelt et al. 1999) assume a strict separation between the semantic/syntactic representation of words and the word forms as discussed above. Evidence for the two-level representations of words comes from tip-of-the-tongue experiments and picture-word interference experiments. Some of these studies are first summarized below, followed by a summary of a study that relate to lexical access in Japanese. (There is also evidence against the serial discrete nature of information flow, see Meyer and Belke 2007 and Vigliocco and Hartsuiker 2002).

When speakers are at the tip-of-the-tongue state (i.e., failing to retrieve the exact forms of words that they know exist to express their meanings), they can report the grammatical properties of their target words above chance level. They can report that the word they were searching for is a count noun or a mass noun (Vigliocco et al. 1999), a masculine noun or feminine noun (Vigliocco, Antonini and Garrett 1997), or in the case of Japanese, adjectives or adjectival nouns (Iwasaki, Vigliocco and Garrett 1998). These findings provide evidence that there are abstract representations of words that contain the words' grammatical information in the absence of their phonological forms.

Utilizing the picture-word interference paradigm, Schriefers, Meyer, and Levelt (1990) provided evidence for serial two-stage word retrieval. In their experiment, the Dutch-speaking participants named pictures while they heard interfering words via headphones (which they had to ignore). In the semantic condition, for example, when a picture of a clock was presented, and thus the participants needed to say the Dutch word *klok*, they heard the semantically related word *horloge* 'watch'. In the phonological condition, they heard a phonologically related word *klos* 'chock'. The timing of the auditory interfering stimuli was manipulated; it was presented at

150 milliseconds before the presentation of the picture, at the same time, and 150 ms after. They found that the semantically related words interfered with the naming at the early stage of lexical access (when auditory stimuli were presented at 150 ms before the presentation of the picture), and the phonologically related words facilitated it later. They interpreted the findings as evidence for the serial activation of semantic representation of words followed by the phonological representation of the word. But these earlier studies which provided evidence for two-stage lexical access were mostly limited to the investigation of nouns, and early picture-word interference experiments did not examine the activation of syntactic properties of the target words either.

Using the picture-word interference paradigm, Iwasaki et al. (2008) examined the grammatical class effect as well as semantic interference in verb retrieval in Japanese, following an earlier study in Italian (Vigliocco, Vinson and Siri 2005). In the Italian study, the verb distracters (e.g., *pettinare* 'to comb'), compared to noun distracters which have action meanings such as *calcio* 'the kick', delayed the naming of target action pictures (e.g., *starnutire* 'to sneeze') when the participants had to produce the inflected form of the verbs (the equivalents of sentence fragments with non-overt subjects in Italian), but not when they named the pictures in single words in citation form. This was taken to indicate the activation of the grammatical class of the target words, which competed with the distracter words for the slot in the syntactic frame being built.

In Iwasaki et al.'s (2008) third experiment, the participants (N = 64) named action pictures either in a single word in citation form (e.g., *oyogu* 'swim' when a picture depicting someone swimming was presented) or in a sentence (e.g., *otoko ga oyoide iru* 'the man is swimming'). The participants read aloud the word presented prior to the picture (e.g., *otoko* 'man') before naming the picture in a phrase to complete the sentence. At the same time as the picture was displayed, the participants heard a distracter word in one of four conditions: semantically similar verbs (*moguru* 'dive' for the swimming picture), semantically dissimilar verbs (*somuku* 'disobey'), semantically dissimilar verbal nouns (*hoyoo* 'recuperate') and semantically dissimilar nouns (*byoobu* 'screen'). A semantic interference effect similar to previous studies was found; it took longer to name a target picture when the distracter word was a semantically similar verb. However, there was no grammatical class effects; participants named the pictures just as fast when the distracters were semantically dissimilar, regardless of the distracter words' grammatical class, and regardless of whether the participants were producing single words or sentences, in contrast to the results in the Italian study.

The presence of the grammatical effect in Italian was attributed to the pressing need to produce the verb that agrees with the subject in gender and number, for which activation of grammatical class information is required early. The absence of a grammatical class effect in the Japanese study was attributed to the relative independence between the subject and verb in sentence production in Japanese. The

result that there was no grammatical class effect may challenge the strictly staged serial lexical access claim (e.g., Levelt 1989), which assumes that the selection of the lemma (which contains grammatical information) necessarily precedes the retrieval of the word form.

Further, Vigliocco and Kita (2006: 806) suggest that examination of mimetic, or sound-symbolic words (e.g., *gorogoro*, which describes the manner in which a heavy object rolls), which are abundant in Japanese, makes it possible to take a new approach to "the issue of whether phonological properties can affect semantic representation and processing". Unlike most other words, the relationship between mimetic words' meanings and forms are not arbitrary. A word's form resembles the word's meanings (what the word refers to). This relationship, iconicity, is found in sound-symbolic words across many languages in the world including sign languages (See Nuckolls 1999; Perniss, Thompson and Vigliocco 2010) but it happens to be very limited in English and European languages and thus its role in language production has rarely been examined so far.

In Japanese, there is consistent phonology-meaning correspondence such as voiced consonants including /d/ and /b/ correspond to large volume (See Hamano 1997). Vigliocco and Kita hypothesize that mimetic words may be produced differently due to the direct co-activation of the word's meaning and phonological representation. They hypothesize, for example, that children learn to correlate conceptual and phonological properties such as big/heavy with [+voicing], and as a result, the activation of one co-activates the other. In case of British sign language, Vinson et al. (2013) found that iconicity facilitated lexical access.[6]

In this section, sentence production processes that are largely agreed upon as well as processes that may be different in Japanese were summarized. The former should help us understand a review of L2 sentence production models below and the latter will help us discuss L2 Japanese sentence production in the next section.

# 3 L2 Sentence production models and L2 Japanese

## 3.1 Overview

Like L1 research, compared to L2 processing/comprehension, research on L2 sentence production from psycholinguistic perspectives is very young; Poulisse (1997: 221) stated that "the development of models of second language production has just

6 Further, iconicity at the sentence level may also play an important role both in language production and comprehension in L1 and L2. Such iconicity is often found between word order and temporal sequences in which entities are involved in the event (see, for example, Tai 1985 for iconicity in Chinese word order; O'Grady, Yamashita and Lee 2005 for L2 learners' preference for isomorphic mapping of word order and event).

begun" in the mid-1990s. There are very few studies on L2 Japanese production from psycholinguistic perspectives which consider the processes we just reviewed above. In this section, models proposed to account for L2 sentence production are reviewed and L2 behavioral experimental studies that were conducted to revise the models are discussed. Implications for L2 Japanese research will also be briefly discussed.

A very important aspect of the now established understanding of bilingual processing/comprehension should be noted. That is, bilingual speakers are different from monolingual speakers in that their linguistic representations and processes are not a combination of monolingual speakers of the two languages they speak. Rather than having two separate systems, their language representations and processes both in their native language and in their L2 are influenced by both of their languages. Their processing "appear to be 'in between' the individual's two codes" (Hernández, Fernández and Aznar-Besé 2007: 371).

Such findings corroborate the current understanding of second language learners' knowledge in studies in second language acquisition (SLA) (e.g., Cook 1992). This is important to keep in mind when we consider L2 speakers' sentence production as well. Note also many researchers use the term "bilingual" to refer to both "balanced bilinguals" who are proficient in two languages, and unbalanced bilinguals (L2 learners). In this chapter, the term is also used inclusively to refer to both balanced bilinguals and any speakers who use two languages unless otherwise noted.

Since the 1990s, several models have been proposed for bilingual sentence production based on L1 sentence production research; the seminal model is the adaptation of Levelt's (1989) model by de Bot (1992). But until recently most empirical studies investigating bilingual production were on lexical representations and access. Studies on L2 structure building research from psycholinguistic perspectives were very limited. Among the studies was a study by Nicol, Teller and Greth (2001) who examined the distributivity effect discussed above on subject-verb agreement among English-Spanish balanced bilinguals and L2 Spanish learners (they found the effect only among bilingual speakers). Recently, however, the syntactic/structural priming experiments (Bock 1986) have been used to shed light on the relationship between the structure-building systems of the two languages among bilinguals.

First, de Bot's (1992) seminal adaptation of Levelt's model is described with some notes to update the model based on more recent studies.

## 3.2 Bilingual sentence production models adapting Levelt's model

Keeping as much of the L1 model proposed by Levelt (1989) as possible and adapting it only when necessary to account for L2 empirical evidence, de Bot (1992) proposed a bilingual sentence production model. Figure 2 is a depiction of the current

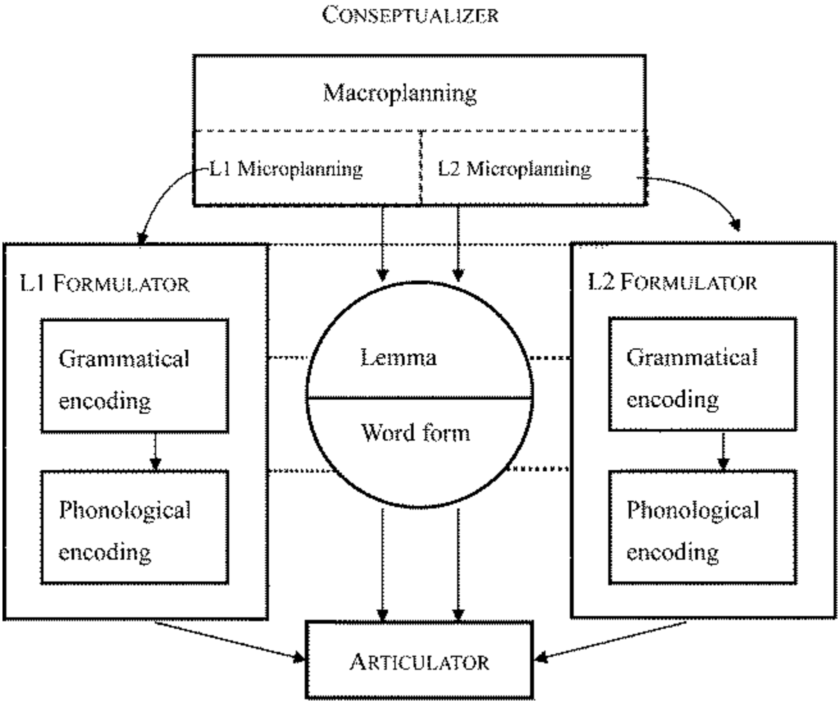



**Figure 2:** De Bot's (1992) adaptation of Levelt's (1989) model depicted by the current author

author's understanding of his adapted model. Some lines are added for clarity, e.g., lines from conceptualizer to the lexicon that must be omitted in Levelt's original figure in 1989 (Levelt 1989: 9). Levelt's original figure depicted monitoring and the relationship between production and comprehension as well, but only the production components, which de Bot adapted, are depicted here.

Of the four major components, de Bot suggests that the formulators may be language-specific especially if the languages are typologically distant. The conceptualizer, the mental lexicon, and the articulator are shared between the languages. Below, de Bot's assumptions and proposals for conceptualizer, formulator(s) and mental lexicon are summarized. (Phonological encoding and the mechanism of the articulator are beyond the scope of this chapter.)

De Bot (1992: 8) suggests the possibility that the first of the two processes in the conceputulizer, namely macroplanning (i.e., planning of how to convey the intended message with the consideration of situational knowledge and discourse), is not language-specific but that microplanning should be language-specific. During microplanning, preverbal messages that include conceptual distinctions required by the lexicalization patterns of the languages are prepared. For example, spatial reference information required to distinguish between *here* and *there* (or *this* and *that*) in

English differs from information required in the three-way distinction in Spanish, *aquí/ahí/allí* (proximal/medial/distal), and the sufficient conceptual information needs to be in the preverbal information.

Somewhat contradictorily, as pointed out by Poulisse (1997, 1999) and Poulisse and Bongaerts (1994), de Bot (1992) assumes that bilingual speakers simultaneously encode the intended message in two languages and produce two speech plans, which accounts, for example, for fluent code-switching. Thus in the model depicted in Figure 2, the two arrows to the two formulators reflect simultaneous information flow for two speech plans. However, this simultaneous preparation of two plans seems to be abandoned in other studies (e.g., Poulisse 1997, 1999). In a more updated model, then, the arrows can be interpreted as two possible paths of processes to choose from.

De Bot suggests that whether there should be two separate formulators or not depends on the linguistic distance between the two languages and on the level of bilingual speakers' L2 proficiency. A shared system may be used by bilingual speakers whose languages are very similar to each other, but separate systems may be required for unrelated languages. Also, beginning L2 learners may use their L1 system to speak in the L2. Poulisse (1999: 169), who examined Dutch-English bilinguals' slips of the tongue, suggests that when two languages are typologically similar, such as in the case of Dutch and English, "L2 learners may accidentally follow L1 instead of L2 syntactic encoding procedures".

With regard to mental lexicon, there are L1 and L2 subsets networked in the lexicon. Adopting an approach with "activation spreading" between elements, such as the model proposed by Dell (1986), de Bot suggests that the items that are in the subset of the chosen language are selected for articulation.

Based on the examination of Dutch-English bilinguals' unintentional use of L1, Poulisse and Bongaerts (1994) support de Bot's assumption of shared mental lexicon with spreading activation between the elements, but they argue against the formulation of two speech plans. They assume, instead, that the language choice is added to the preverbal message and that each lemma is tagged with a language label. If a preverbal message is specified for L2, then lemmas tagged as L2 receive more activation. They argue that an unintentional use of an L1 Dutch word in L2 English is an error of accessing the L2 lemma due to the fact that the L2 lemma and L1 lemma, which is the L2 word's translation equivalent, are both highly activated.

While many researchers tackled the issue of how bilingual speakers manage to control and select the language they intend to use (See, for example, La Heij 2005), there is now ample evidence that words in both languages receive activation as shown by unintentional switches observed by Poulisse and Bongaerts (1994), for example. According to Costa (2005: 312), current models of bilingual speakers' lexical access "favor the idea that activation from the conceptual system flows to lexical representations of both languages of a bilingual". The question is how the existence/activation of lexical representations in one language affects the selection

processes in the other language, and this may depend on the speaker's L2 proficiency level.

## 3.3 Is syntax shared by two languages? Cross-linguistic syntactic priming

Until recently most empirical studies on L2/bilingual cognitive processes during language production investigated lexical issues (i.e., lexical representation, selection, and retrieval). Since the early 2000s, however, a number of researchers have been investigating bilingual speakers' structure building utilizing syntactic priming (also called "structural priming") experiments. In L1 research, since Bock's (1986) study, it has been demonstrated in studies conducted in many languages (albeit mostly European languages) that speakers tend to reuse structures that they recently used when there are alternative structures to convey the same/similar meanings, such as active/passive or NP-NP/NP-PP dative (e.g., *give a friend a gift* vs. *give a gift to a friend*) sentences. In syntactic priming experiments, participants typically hear or say a sentence in one of the alternative structures (such as a passive sentence) and then produce a target sentence by describing a picture or by recalling a sentence that they heard or read earlier. In the target sentences, the participants are found to reuse the structure they heard or said earlier rather than using the other alternative (such as an active sentence).

The finding from syntactic priming was interpreted as evidence for the abstract representation of syntactic structure because the effect was observed regardless of whether the prime and target sentences shared closed-class words (*for* or *to*) (Bock 1989), or thematic roles of the components (e.g., *by*-agent in passive, *by*-location) in the sentences (Bock and Loebell 1990) as long as the primes share the same structure as the targets such as V-NP-PP. The nature of primed structural representation (hierarchical structures/dominance, or linear order) and the locus of priming (functional, positional level, or merged one level) do not seem entirely clear (see Ferreira and Slevc 2007). Yet, Costa et al. (2007: 538) state that the experimental paradigm appears to be "the most relevant observations for understanding how grammatical encoding proceeds".

Recently, priming effects have been found among bilinguals and across languages (e.g., Loebell and Bock 2003; Meijer and Fox Tree 2003; Shin and Christianson 2009). Based on their study of Spanish-English bilinguals, Hartsuiker, Pickering, and Veltkamp (2004) argue for shared syntax among bilinguals. In their experiment, after English-Spanish bilingual speakers hear the experimental confederates' description of a picture in Spanish in one of four structures (active, passive, intransitive, Object-Verb-Subject sentences), they examine whether a picture they have in front of them corresponds to what was described, and press yes/no button. They then describe a picture on the next card in their card stack in English. The participants

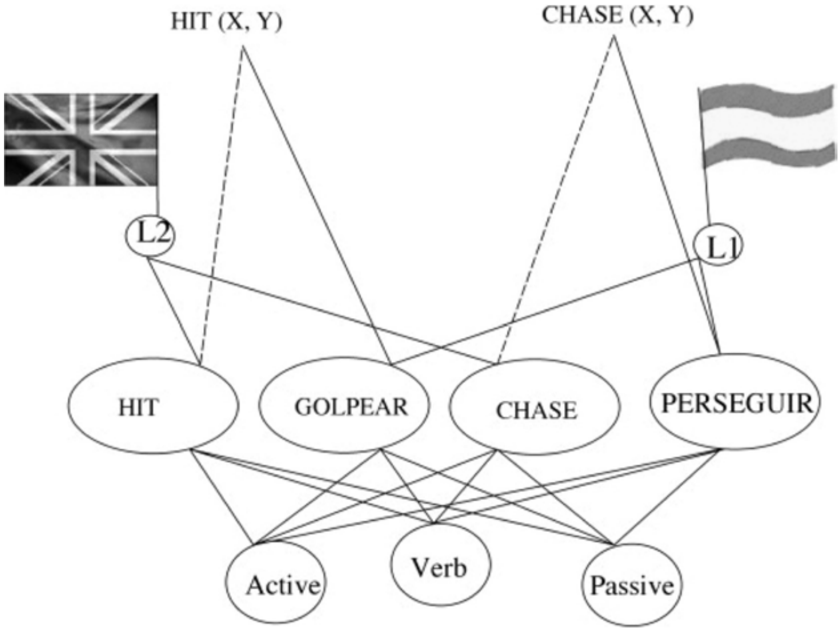



**Figure 3:** Example of Spanish-English bilinguals' lexical entries for 'to chase' and 'to hit' (Hartsuiker, Pickering and Veltkamp 2004)

tended to produce English passive sentences after having heard a Spanish passive sentence, providing evidence for cross-linguistic syntactic priming. Hartsuiker et al. adopt lexically driven grammatical encoding (Pickering and Branigan 1998) and propose the bilingual lexical representation shown in Figure 3 (Hartsuiker et al. 2004: 413). In their model, syntactic information needed for grammatical encoding is represented at lemma nodes, which are linked to combinatorial nodes (encoding combinatorial information). For example, the verb *chase*, which can be used either as part of passive or active sentences, are associated with two nodes. These combinatorial nodes are shared between different verb lemmas (e.g., *hit*, *chase*) across different languages (Spanish translation equivalents *golpear, perseguir*). These lemmas also share the same conceptual nodes, HIT (X, Y) and CHASE (X, Y). When a link between a lexical representation and a combinatorial node is activated, then its residual activation of the syntactic representation leads to syntactic priming.

Hartsuiker and Pickering (2008) evaluated Hartsuker et al.'s (2004) model along with de Bot's (1992) adaptation of Levelt's model and another model proposed on the basis of evidence of neural studies (Ullman 2001) by testing different predictions made in these studies. They conclude that the lexically driven model proposed by Hartsuiker et al. can better account for the empirical evidence that is currently available.

Hartsuiker et al. (2004) suggested that the syntax of a particular construction is shared between languages of bilinguals if it is formed similarly in the two languages because speakers "strive towards an economy of representation" (Hartsuiker 2013: 739). Their suggestion was to account for lack of priming for passive constructions among English-German bilinguals reported by Loebell and Bock (2003). Loebell and Bock found cross-linguistic syntactic priming for dative constructions, which is comparable in English and German, but not for the passive construction. With evidence emerging across other languages, most notably between English and Korean (Shin and Christianson 2009), Hartsuiker and Pickering (2008: 485) argue that Hartsuiker et al. (2004) "predict no difference between cross-linguistic priming in closely related languages (e.g., Dutch and English) or very distant languages (e.g. Korean and English), as long as the languages have similar syntactic rules."

In earlier studies, such as Bernolet, Hartsuiker and Pickering (2007), word order overlap between L1 and L2 was considered to be a requirement of structural similarity for cross-linguistic priming to occur. They examined whether there was cross-linguistic priming among Dutch-English and Dutch-German bilinguals for relative clause construction, and they only found priming between German and Dutch, which share the same word order. This was taken as evidence that bilingual speakers may not have shared representations of structures differing in word orders in two languages.

Shin and Christianson (2008), however, argue that the syntactic priming they observed among English-Korean bilinguals suggest shared abstract syntactic representation at the functional level processes that are independent of word order. Using a recall paradigm similar to Meijer and Fox Tree's (2003) experiments, they examined whether Korean dative structure primed English dative structure. The participants heard an English target sentence, which was either a double-object (NP-NP) or prepositional (NP-PP) dative. They then heard a Korean prime sentence in one of the three dative structures shown below: (14) the canonical-order postpositional dative, (15) the double-object dative, and (16) the scrambled postpositional dative. After hearing the Korean sentence and completing a distraction task[7], the participants recalled the English sentence they heard at the beginning of the sequence.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (14) | *Mary* | *ka* | *John* | *eykey* | *chayk* | *ul* | *cwuessta.* |
| | | NOM | | to | book | ACC | gave |
| | 'Mary gave a book to John.' | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (15) | *Mary* | *ka* | *John* | *ul* | *chayk* | *ul* | *cwuessta.* |
| | | NOM | | ACC | book | ACC | gave |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (16) | *Mary* | *ka* | *chayk* | *ul* | *John* | *eykey* | *cwuessta.* |
| | | NOM | book | ACC | | to | gave |

7 In recall tasks, some distraction tasks are typically given to participants prior to recalling the target sentences. In this case, the participants judged whether the Korean word shown on the computer screen was in the Korean sentence they heard.

The participants produced more prepositional constructions in English after the prime sentence type (14), which is considered as the equivalent of English prepositional dative with different word order. Thus Shin and Christianson conclude that there is shared bilingual processing occurring at functional level, supporting the two-stage process.

With emerging evidence, Hartsuiker (2013) also no longer considers word order overlap to be a requirement for cross-linguistic priming. He cites, for example, Bernolet, Hartsuiker and Pickering (2009). They observed priming between English passive (by-phrase final) and Dutch passives (verb-final), differing in word order. Adopting Hartsuker et al.'s (2004) model, Hartsuiker (2013) argues that priming takes place at the level of combinatory units because bilingual speakers search for correspondence between two languages (e.g., "active" and "passive" in Figure 3 above). Further, based on Bernolet et al.s' (2009) findings, he argues that priming also takes place at a level where that is concerned with thematic role ordering.

The extent of the priming effect was found to depend on the L2 proficiency of bilinguals. Bernolet, Hartsuker and Pickering (2013) examined cross-linguistic priming among English-Dutch bilinguals, utilizing the genitive "of" and "-'s" constructions. Though both languages have similar "of" construction, the realization of "-'s" is different between the two languages (e.g., *the nun's hat* in English; *de non haar hoed* "lit. the nun her hat"). They found that the cross-linguistic priming effect linearly increased with L2 proficiency. Hartsuiker (2013: 738) thus suggests that L2 learners start with separate syntax for their L2 (specifically, separate combinatory nodes, in Hartsuker et al.'s (2004) model), and later "collapse" them with their L1 representation, which results in shared syntax.

## 3.4 What about structure-building in L2 Japanese?

There are very few studies both on L2 Japanese lexical access and structure building from psycholinguistic perspectives, though there are emerging studies on how Japanese speakers describe motion events, which sheds lights on how the conceptualizer and formulator may be related. This will be discussed in Section 4.

There appear to be no syntactic priming experiments conducted with bilinguals who speak Japanese as one of their languages, but syntactic priming experiments in Japanese have been conducted, and priming patterns observed seem to be different from syntactic priming in other languages. Because this has bearing on the potential significance of bilingual syntactic priming experiments, what has been found about syntactic priming in Japanese is briefly summarized here first.

Unlike Shin and Christianson's (2009) assumption that English NP-PP dative and Korean PP-NP dative share the same structure, two alternative Japanese dative structures are generally both considered to have the NP-NP structure. Yamashita, Chang and Hirose (2003) tested whether NP-DAT NP-ACC dative, (18) below, and NP-LOC

NP-ACC, which is (19), could prime the same NP-DAT NP-ACC dative structure, using an immediate recall task (Potter and Lombardi 1998). Japanese-speaking participants read the target sentence (ACC-DAT dative construction, Agent-Theme-Beneficiary-Verb like (17)) presented rapidly phrase by phrase on the computer screen, and recalled it after a distraction task. The prime sentence that preceded the recall trial was one of three types: (18) DAT-ACC, (19) LOC-ACC, and (20) ACC-DAT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (17) | *Obaasan* | *wa* | *syakkin* | *o* | *komeya* | *ni* | *haratta.* |
| | grandmother | TOP | debt | ACC | rice.shop | DAT | paid |

‘The grandmother paid the debt to the rice shop.’

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (18) | *Syatyoo* | *wa* | *roozin* | *hoomu* | *ni* | *wagonsya* | *o* | *kizoo-sita.* |
| | CEO | TOP | elderly | home | DAT | wagon | ACC | presented |

‘The CEO presented the wagon to the retirement home.’

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (19) | *Syatyoo* | *wa* | *roozin* | *hoomu* | *ni* | *wagonsya* | *o* | *tyuusya-sita.* |
| | CEO | TOP | elderly | home | LOC | wagon | ACC | parked |

‘The CEO parked the wagon at the retirement home.’

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (20) | *Syatyoo* | *wa* | *wagonsya* | *o* | *roozin* | *hoomu* | *ni* | *kizoo-sita.* |
| | CEO | TOP | wagon | ACC | elderly | home | DAT | presented |

‘The CEO presented the wagon to the retirement home.’

The participants swapped the ACC-DAT sequence of the target sentence to DAT-ACC significantly more often (21%) when the prime sentence was type (18) DAT-ACC than when the prime was (19) or (20) (9% and 11%, respectively) despite the fact that both prime sentences (18) and (20) are analyzed as NP-NP like the target sentence (17). Moreover, superficially similar (19) PP-NP (LOC-ACC, NP-*ni* NP-*o*) structures did not prime DAT-ACC structures. The Recipient-*ni* Patient-*o* sequence appears to be treated differently from the Location-*ni* Patient-*o* sequence. Yamashita et al. suggest the possibility that what is being primed is the mapping between meaning (i.e., the semantic roles of the NPs) and the grammatical functions.

This diverges from the finding in English (Bock and Loebell 1990) that active intransitive sentences with *by*-locative prepositional phrases (e.g., *The 747 was landing by the control tower*) prime passive sentences (e.g., *The 747 was alerted by the control tower*) despite the differences in the conceptual features of the prepositional phrases, but is compatible with findings by Chang, Bock and Goldberg (2003). They conducted experiments using two types of English sentences with the same structure (V NP PP) varying the order of the thematic roles of the arguments in their syntactic priming experiments. They found that thematic role arrays mattered in syntactic priming. Theme-Location sentences like *The maid rubbed polish onto the table* primed Theme-Location sentences such as *The farmer heaped straw onto the wagon*

and Location-Theme sentences like *The maid rubbed the table with polish* primed Location-Theme sentences such as *The farmer heaped the wagon with straw*. Chang et al. suggest the possibility that there may be a mechanism that associates thematic role arrays with structure configuration like the argument-structure construction (Goldberg 1995) in meaning-form mapping during sentence production.

Yamashita et al. found that a specific array of thematic roles of the prime (i.e., Recipient-Patient sequence as in Example (18)) primes Recipient-Patient sentences, swapping Patient-Recipient in the target sentence (17). This finding supports the possibility that the NP thematic array plays a significant role in Japanese syntactic priming. Other studies on Japanese also revealed the influence of meaning-related factors such as animacy (Tanaka et al. 2011 discussed above) and givenness (i.e., whether an entity/participant is already given/mentioned in the discourse) (Ferreira and Yoshita 2003) on Japanese speakers' choice of alternative word orders. Together, there is indication that functional assignments and constituent assembly make reference to conceptual features more so than assumed before, suggesting the close relationship between the conceptualizer and the working of the formulator, involving both functional and positional processes.

Would English-Japanese bilinguals be predicted to show shared syntax of the two languages by demonstrating cross-linguistic syntactic priming? On the one hand, given the differences found between English and Japanese language processing during sentence production, bilingual speakers of Japanese and English are unlikely to use the shared syntax to process the two languages, and cross-linguistic syntactic priming may not be observed among English-Japanese bilinguals.

On the other hand, however, it is plausible that bilingual speakers' linguistic representations and processing are different from those of monolingual English speakers and those of Japanese monolingual speakers, similarly to what has been found in other studies in SLA (see Section 5 below). Thus, for example, rather than English prepositional datives being intrinsically equivalent to Korean postpositional dative, bilinguals tested by Shin and Christianson (2009) may treat them as equivalent implicitly or explicitly in processing their L2 (in this case English) by searching for correspondence (Hartsuiker 2013). If this is the case, it would not be surprising if English-Japanese bilinguals develop shared syntax for English and Japanese for structures that are linguistically analyzed as different: e.g., English NP-PP datives and NP-ACC NP-DAT Japanese datives so long as the bilinguals themselves perceive them as equivalents. Given that there are more empirical studies on L1 Japanese language production than other non-European languages, research on cross-linguistic syntactic priming among bilinguals who have Japanese as one of their languages is likely to contribute to our understanding of the bilingual language processes.

L2 Japanese sentence production processes are clearly under-researched despite the fact that Japanese is one of the most studied non-European L2s in SLA. There are a number of studies on L2 acquisition of structures in Japanese, such as relative clauses and case-markers, using the production data (spoken data, written data)

(see Mori and Mori 2011 for a review), and some studies are relevant to L2 Japanese sentence production (See Iwasaki 2003; Iwasaki 2006b). However, because most of these studies do not adopt experimental paradigms that are used to investigate L1 sentence production processes, it is difficult to relate them to previous studies in the field.

One area in which there is currently a critical mass of research on cognitive processes involved in speaking with L1 and L2 Japanese speakers is conceptualizing motion events for speaking and communicating.

# 4 Thinking-for-speaking

## 4.1 Talking about motion: Language-specific package of manner and path

Though the strong version of linguistic relativity that claims that language determines worldview or habitual thought (e.g., Whorf 1956) has largely been disputed, a weaker formulation, the thinking-for-speaking hypothesis (Slobin 1996) has drawn a good deal of attention and has been studied in recent years. Languages differ in what is obligatorily coded grammatically or lexically. Thus, Slobin (1996: 15) argues, "There is a special kind of thinking that is intimately tied to language – namely, the thinking that is carried out on-line, in the process of speaking" because different languages direct us to pay attention to differing dimensions of experience.

In particular, languages are found to differ greatly in lexicalization patterns to describe motion events (i.e., in the ways semantic elements of motion events are mapped to lexical units or grammatical categories) (Talmy 1985, 2000) and children learning their languages show sensitivity to language-specific patterns (Allen et al. 2007; Choi and Bowerman 1991). A motion event is very complex; the rich information needs to be organized and packaged so that it can be verbalized in a given language.

According to Talmy (e.g., 2000), a basic motion event consists of four internal components: Figure (the moving or stationary object); Ground (the object in relation to which the Figure moves or is located); Path (the path followed by or the site occupied by the Figure); Motion (the presence of motion or locatedness in the event). In addition, there are two external components: Manner of motion and Cause of the occurrence of event. Based on the ways they express Path and Manner of motion, many languages can be classified either as "Satellite-framed languages" (S-languages) or "Verb-framed languages" (V-languages). English as well as many other Indo-European languages, except Romance languages, are in the former category, and Japanese as well as Korean, Turkish, and Romance languages are in the latter category. In the English expression in (21), the verb *roll* indicates both Motion

and Manner, and the particle (satellite) *down* expresses Path. In contrast, in the Japanese sentence in (22), the main verb *kudaru* 'descend' expresses Motion and Path, and Manner is encoded as a gerund of a verb *korogat-te* 'rolling'.

(21) *He rolls down the hill.*

(22) *Korogat-te saka o kudaru*
roll-GER slope ACC descend
'(He) descends the slope, as (he) rolls'

Slobin (1996) reports, for example, that English and other S-languages use a variety of manner-of-motion verbs and thus Manner is more salient among S-language speakers than in V-language speakers. S-language users also exhibit higher degrees of elaboration of Manner. He suggests that children learn particular ways of thinking-for-speaking for their L1 and learn to pay attention to the dimensions of experience that need to be encoded in their language.

## 4.2 Talking and gesturing about motion events

Motion is not expressed by language alone; it is concurrently encoded in gesture. Kita and Özyürek (2003) proposed a hypothesis to account for motion event descriptions that involves both speaking and gesturing. In their Interface Hypothesis, there exists spacio-motoric imagery from which the gesture originates, and this imagery is shaped both by the language requirement (lexicalization patterns and processing units for sentence production) and by the spatio-motoric properties of the referent, which is not always expressible in the language. Their hypothesis predicts that the gestural expressions of events vary across speakers of different languages in accordance with the linguistic expressions available to describe the events. It also predicts the use of gestural expressions that are independent of language. They supported their hypothesis by comparing descriptions of motion events among speakers of English (S-language) and Japanese and Turkish (both V-languages). Kita and Özyürek examined descriptions of two events in which both Manner and Path are salient. A Swing event (i.e., swinging on a rope across two buildings) and a Rolling event (i.e., rolling down on a slope into a bowling alley) were selected from an animated cartoon (the Tweety Bird).

Unlike English, neither Turkish nor Japanese have a single verb depicting an arc-shaped trajectory for swinging. In describing the Swing event, almost all English speakers used the verb "swing", which entails an arc-shaped Path, but Japanese and Turkish speakers encoded the event without lexically encoding the arc-shaped Path. In accordance with the speech, while English speakers often used arc-shaped gesture, Japanese and Turkish speakers tended to use more straight-line gestures.

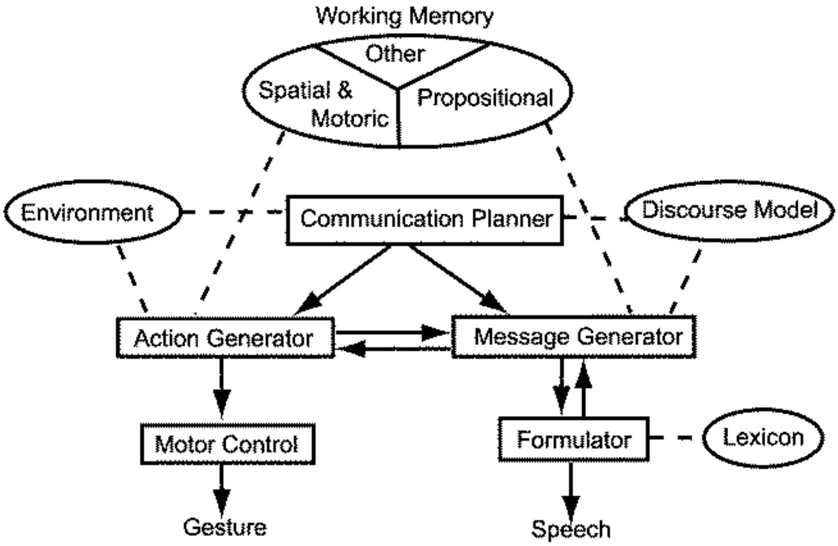



**Figure 4:** Kita and Özyürek's (2003) model of speech and gesture production

But all language groups used a leftward gesture to depict the leftward movement, which was never encoded verbally.

In describing the Rolling event, Kita and Özyürek (2003) found, as expected from the linguistic differences, that English speakers encoded both Manner (i.e., rolling) and Path (e.g., along the road, down the slope) in the same clauses (i.e., in one processing unit) like (21) above while most Japanese speakers used two clauses to encode Manner and Path as in example (22) above. Though Manner can be expressed by using mimetic words, according to Kita and Özyürek (2003: 27), such mimetic expression is "typically intonationally separated from the Trajectory expression". English speakers primarily used Manner-Path conflating gestures, but Japanese speakers used Manner-only and Path-only gestures in addition to Manner-Path conflating gestures, reflecting how the information is packaged in each language.

Allen et al. (2007) also compared English, Japanese and Turkish speakers' descriptions of motion events in which both Manner and Path are salient. They studied both adult speakers and 3-year old children. Similarly to Kita and Özyürek (2003), they found that both English speaking adults and children encoded Motion and Path in single clauses, and adult Japanese speakers used multiple clauses where one of them was typically a subordinate clause) – although Japanese children often encoded Manner and Path in single clauses.

Kita and Özyürek (2003) present their model of speaking and gesturing, building upon Levelt's (1989) model, shown in Figure 4 (Kita and Özyürek 2003: 28). The main characteristic of the model is that the conceptualizer is split into two. One is

the Communication Planner, which not only deals with macroplanning in Levelt's model but also determines which modalities of expression (i.e., gesture or speech) should be used. The other is the Message Generator, whose function is similar to Levelt's microplanning. Importantly, in order to account for coordination between linguistic expressions and gestures, there is on-line feedback from the Formulator to the Action Generator via the Message Generator. The information flow between the Action Generator and the Message Generator, and between the Message Generator and the Formulator is bi-directional. This last aspect of the model is significantly different from Levelt's. The online bi-directional information between the Message Generator and the Formulator was supported by further evidence (Allen et al. 2007) which demonstrated cross-linguistic differences in gesture were linked to the syntactic structures of the co-expressive language in English, Japanese, and Turkish (see also Kita 2010 for further specification and development of the model).

If there is on-going bidirectional interaction between the Formulator and the Message Generator and between the Message Generator and the Action Generator, a number of questions arise for bilingual speakers' motion event descriptions. For instance, given the evidence that we reviewed in Section 4 supporting simultaneous activation of bilinguals' two languages, are bilingual speakers' ways of syntactic packaging simultaneously influenced by two languages? Which language's syntactic packaging do their gestures reflect? Do L2 learners shift their L1 patterns of conceptualization (thinking-for-speaking) of motion events to L2 patterns when speaking L2? Some of these questions have been investigated.

# 5 Thinking for speaking and gesturing in L2

## 5.1 Talking and gesturing about motion in L2

Typological differences in motion descriptions in Talmy's framework allow SLA researchers to investigate cross-linguistic influence in form-meaning mapping. Cadierno (2008: 158) states that two SLA questions in terms of Slobin's thinking-for-speaking hypothesis are: "how and to what extent do adult L2 learners adapt to their thinking-for-speaking in an L2 that is typologically different from their L1, and how does the adaptation of this type of learner compare to that followed by learners whose L1 and L2 share the same typological patterns?" With regard to motion description (i.e., Manner and Path) by speech and gesture, the first question is tackled by research comparing L1 Korean (a V-language, like Japanese) speakers and English-Korean bilingual speakers (Choi and Lantolf 2008) and research examining L1 Japanese-L2 English bilingual speakers (Brown and Gullberg 2008; 2011; 2012). In addition, Yoshioka and Kellerman (2006) studied the description of Ground among L1 Dutch (S-language, like English) speakers, L1 Japanese speakers, and L1 Dutch-L2 Japanese bilingual speakers.

Choi and Lantolf studied 4 nearly balanced bilingual speakers (2 L1 English speakers who are highly proficient in L2 Korean and 2 L1 Korean speakers who are highly proficient in L2 English) in order to find whether bilingual speakers shift their L1 thinking-for-speaking patterns to L2 patterns. They also collected L1 Korean data to understand L1 Korean patterns, which have not been extensively studied, unlike L1 English patterns that are widely reported. Using scenes from the same animated cartoon that Kita and Özyürek (2003) used, they compared bilingual speakers' speech and gesture to monolingual Korean speakers' and monolingual English speakers' speech and gesture. They found that in terms of Path-only gestures, the bilinguals' L2 gestures approximated L1 monolingual speakers' Path-only gestures in that they co-occurred with linguistic elements that expressed Path, Ground, or both in the respective language.

For Manner, however, both L2 English and L2 Korean speakers retained their L1 patterns of gestures. For descriptions of the rolling event in which Manner is salient, neither L2 English speakers showed sensitivity to Manner. One did not use any manner verbs; the other used manner verbs but without co-speech manner gestures unlike monolingual English speakers who tend to produce Manner-Path conflated gestures. Instead, L2 Korean speakers were sensitive to Manner when describing scenes in which manner of motion is not salient. They were unable to access the low-frequency Korean verb *kwull-e* 'roll', and became dysfluent when they attempted to describe manner of motion. Hence, Choi and Lantolf concluded that thinking-for-speaking to describe manner of motion in L2 requires a conceptual shift and thus is not easily attainable.

Yoshioka and Kellerman (2006) also found that L1 Dutch speakers learning L2 Japanese (with low intermediate proficiency) retained L1 patterns of encoding Ground both in speech and gesture when they describe a story depicted by pictures in *Frog, Where are You?* (Mayer 1969). Unlike L1 Japanese speakers, the L2 Japanese speakers introduced Ground as part of the VP most of the time (e.g., . . . *otokonohito o* <u>*ike*</u> *ni otosimasu* '(the animal) dropped the man into <u>a pond</u>') while L1 Japanese speakers also used independent clauses to introduce Ground (*asai* <u>*kawa mitaina tokoro*</u> *ga aru no* 'there is a shallow <u>river-like place</u>.') Their gesture that co-occurred with the introduction of Ground in speech depicted the action or direction and form of the referent equally often, while L1 Japanese speakers nearly always depicted the outline of the referent. Thus both L2 learners' speech and gesture retained L1 patterns of thinking-for-speaking that are commonly observed among S-language speakers, paying more attention to dynamics of movement rather than static scenes (see Slobin 1996).

## 5.2 Talking and gesturing about motion in L1 Japanese and L2 English

Whereas Choi and Lantolf (2008) and Yoshioka and Kellerman (2006) were examining whether and how L2 learners shift their thinking-for-speaking patterns to L2

patterns, Brown and Gullberg's (2008, 2011, 2012) regard the relationship between L2 learners' L1 and L2 as bidirectional in that each language influences the other.

The bilingual speakers they studied are L1 Japanese speakers whose L2 English proficiency is at the intermediate level. Brown and Gullberg compared these bilingual speakers' speech and gesture to those of monolingual English and monolingual Japanese speakers'. Brown and Gullberg (2008) examined these participants' speech and gesture encoding of Manner and found that both L1 Japanese and L2 English of these bilingual speakers' speech and gesture were between monolingual English and monolingual Japanese speakers'. They produced more spoken descriptions of Manner than monolingual Japanese speakers, but less than monolingual English speakers. They produced less manner fogs (i.e., gestures expressing manner with no manner information in the accompanying speech, a type of gesture typically observed among V-language speakers) than monolingual Japanese speakers, but more than monolingual English speakers. Though monolingual English and monolingual Japanese speakers' performances were significantly different from each other, bilingual speakers' L1 Japanese and L2 English performances were not different from each other.

Brown and Gullberg (2011; 2012) also found that the bilingual speakers were different from both monolingual English and monolingual Japanese speakers in encoding Path of motion in speech. Brown and Gullberg (2011) found, for example, that the bilingual speakers used more Goal Path expressions (adverbial phrases with prepositions or postpositions such as *ni/made, to/until*; Path verbs) in their L1 Japanese and L2 English per clause than both monolingual groups, who did not significantly differ from each other. Brown and Gullberg speculated that this may be because bilinguals employed both the encoding systems preferred in English (using adverbials with *made 'until'; ni 'to'*) and the system preferred in Japanese (using verbs such as *tadorituku* 'arrive', *reach*) when speaking either of the two languages while monolingual speakers had clear preferences.

The bilingual speakers were also significantly different from both monolingual groups in clausal packaging of Manner and Path. Brown and Gullberg (2012) examined the descriptions of four motion events from the same cartoon, including the Rolling and the Swing events. Interestingly, contrary to what Kita and Özyürek (2003) and Allen et al. (2007) found, Brown and Gullberg found that monolingual Japanese speakers encoded Manner and Path in the same clauses nearly as often as monolingual English speakers by encoding both Manner and Path in ways such as (23)–(26) (Brown and Gullberg 2012: 8)[8]. The square brackets indicate the clause, and the elements expressing Manner and Path are underlined.

**8** The notation in glosses and romanization methods used by Brown and Gullberg (2012) were modified for the sake of consistency in the current chapter.

(23)

| | |
|---|---|
| [*korogatte* | *iku*] |
| rolling.GER | go.NONPST |

‘(He) goes rolling.’

(24)

| | | | | |
|---|---|---|---|---|
| [*guruguru* | *gororo* | *to* | *haitte* | *itte*] |
| MIM | MIM | COMP | enter.GER | go.GER |

‘(He) enters going ROLL ROLL.’

(25)

| | | | |
|---|---|---|---|
| [*heya* | *ni* | *tobi-uturoo* | *to*] |
| room | LOC | fly-try.to.move | COMP |

‘(he) tries to fly to the room.’

(26)

| | |
|---|---|
| [*koo* | *yozi-nobotte*] |
| like | clamber-climb.GER |

‘(He) climbs up.’

In the above examples, the packaging of Manner and Path in the single clause is made possible by using the gerundive form of the verb *korogaru* ‘roll’, used similarly to the English participle, combined with a Path verb *iku* ‘go’ in (23), by the use of Manner mimetics combined with Path verbs in (24), by the use of a Manner-Path compound verb *tobi-uturu* ‘fly-move’ and a Path postposition *ni* ‘to’ in (25), and with another Manner-Path compound verb *yozi-noboru* ‘climb up’ in (26).

Interestingly, however, bilingual speakers used more multi-clausal packaging of Manner and Path both in their L1 Japanese and L2 English. Brown and Gullberg speculate, somewhat similarly to their speculation with regard to Goal Path expressions, that “(i) the presence of L2 English causes Manner to be expressed in main verbs in L1 Japanese and that (ii) the presence of L1 Japanese causes Path to be expressed in main verbs in L2 English” (p. 13). As a result the bilingual speakers often employ separate clauses, each packaging Manner and Path in main verbs in each of their languages.

Such interaction of the two languages further supports the simultaneous activation of two language systems both in Lexicon and Formulators (in such models as depicted in Figure 2). Questions that may emerge for those who are interested in examining L2 Japanese are: whether L2 Japanese speakers retain L1 patterns in describing motion (i.e., Manner and Path) similarly to what Yoshioka and Kellerman (2006) found, and whether patterns similar to what Brown and Gullberg (2011, 2012) observed can be observed among L1 English speakers learning L2 Japanese. To my knowledge, there are no published studies on these questions, but the current author’s preliminary analysis of motion descriptions by English speakers learning L2 Japanese suggest that while shifting to L2 patterns is not clearly observed, bi-directional influence is apparent.

## 5.3 Talking and gesturing about motion in L1 English and L2 Japanese

Iwasaki (2013) examined the Swing event and the Rolling event descriptions by 13 L1 English speakers learning L2 Japanese (whose proficiency was low intermediate to advanced) as well as functionally monolingual Japanese speakers. The participants described the same animation clips selected from Tweety Bird as the ones used in the previous studies discussed above. Their shift to L2 Japanese was not clearly observed. Instead, both their L1 English and L2 Japanese speech and gesture produced by speakers who are relatively proficient in L2 Japanese showed patterns that incorporated both oft-reported English and Japanese patterns as discussed below. Only a few of them managed to encode Manner in speech (and/or gesture) in L2 but when they did, the L2 Manner description appeared to reflect L1 and L2 patterns at the same time.

In describing the Swing event, monolingual Japanese speakers used verbs such as *iku* ‘go’ as in (27) and *tobu* ‘fly’ such as shown in (28).

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| (27) | *Huriko* | *no* | *yooni* | *site* | *sono* | *tonari* | *no* | *tatemono* | *made* | *koo* | <u>*itte*</u> |
| | pendulum | GEN | like | do | that | next | GEN | building | as.far.as | this.way | go |

‘He goes to the next building this way like Tarzan.’

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (28) | *Koo* | *taazan* | *mitaini* | <u>*tonde*</u> | *iku* | *n* | *desu* | *kedo* |
| | this.way | Tarzan | like | fly.GER | go | NMLZ | is | but |

‘(It is that) (he) goes flying like Tarzan like this.’

The trajectory depicted in their gestures were arc-only, straight-only, or both, nearly equally often, similarly to what Kita and Özyürek (2003) found. Kita and Özyürek found that L1 English speakers used the verb *swing* and arc-only gesture most of the time. Of 9 L2 Japanese speakers who described the Swing event, 6 used the verb *swing*, but 3 used *fly* or *go* in L1 English. In their L2 Japanese description, some of them used *iku* or *tobu* in L2 Japanese, but others appeared to be searching for the Japanese translation equivalent of the English *swing* and 2 borrowed the English word *swing* and said *suingu suru*. When speaking L1 English, many of them used the arc-shaped gesture, but two used straight-only gesture. When speaking L2 Japanese, only 3 speakers who encoded the Swing event by gesture used an arc-only gesture.

Encoding Manner when describing the Rolling event in Japanese was found to be very challenging for L2 Japanese speakers. In English, most of the bilingual speakers encoded Manner and Path of the Rolling event in single clauses, but only 3 of the 13 bilingual speakers managed to encode Manner in L2 Japanese. They used the verb *korogaru* and/or mimetic words, but the ways one of them used the mimetic

words is clearly affected by their L1 pattern. The excerpts (29)–(32) below are English and Japanese descriptions produced by one of the L2 Japanese speakers.

(29) *then he rolls down a hill, all the way into, into this bowling alley*
*and then he crashes into the bowling alley,*

The phrase *rolls down a hill all the way into this bowling alley* tightly encodes Manner and Path by the use of the manner verb *roll* and the Path adverbial phrase using a preposition *into*, a typical English pattern. There was no co-speech gesture with *he rolls down a hill*, but Path gesture was used with the expression *all the way into*. The speaker extended his right-hand fingers, moved downward to the left in a couple of stepwise movements.

In the excerpt (30)–(32), the English translation given reflects what the bilingual speaker might have meant by his use of mimetic expressions, *korokoro suru, koron to suru*, and *doon*.

| | | | | | | |
|---|---|---|---|---|---|---|
| (30) | *neko* | *ga* | *anoo* | *maa* | *korokoro* | *si-te,* |
| | cat | NOM | uh | well | MIM | do-GER |

‘The cat, uh, well rolled,’

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (31) | *de,* | *sono* | *saka* | *no* | *sita* | *ni* | *booringuzyoo* | *ga* | *arimasita.* |
| | and | that | slope | GEN | below | LOC | bowling.alley | NOM | existed |

‘and there was a bowling alley beneath’

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (32) | *de,* | *neko* | *ga* | *booringuzyoo* | *ni* | *koron* | *to* | *si-te* | *doon.* |
| | and | cat | NOM | bowling.alley | LOC | MIM(roll) | COMP | do-GER | MIM |

‘the cat rolled down to the bowling alley and crashed.’

This speaker primarily used the Path-only gesture when saying (30) and (32) with very subtle circular movement accompanying (32), and a gesture depicting a box shape with (31), suggesting an L1-type Ground depiction described by Yoshioka and Kellerman (2006). Though mimetic words *korokoro* and *koron* depict Manner of rolling, typically the mimetic verbs *korokoro suru* and *koron to suru* do not.[9] Yet, this speaker opted for a creative use of these words as if they corresponded to the English expression *roll down* [Manner + Path].

Most monolingual Japanese speakers used mimetic adverbs, whose semantic representations are argued to diverge from non-mimetic words (Kita 1997; 2001). In

**9** The expression *korokoro-suru* is usually used in the *–ta* or *–te iru* forms as pre-nominal modifiers to describe the state of being chubby, or plump, rather than referring to rolling motion (see, for example, Kakehi, Tamori and Schourup 1996).

the model depicted in Figure 4, the mimetic adverbs may be more strongly associated with a spatial and motoric component than a propositional component as Kita found that mimetic adverbs often co-occurred with iconic gestures. This may imply that L1 monolingual and L2 Japanese speakers' representations and processing of mimetic words may be fundamentally different.

# 6 Conclusion

As mentioned in the introduction of this chapter, production research from psycholinguistic perspectives on L2 Japanese or on bilinguals who speak Japanese as one of their languages is very limited. In recent years, research on language production in L1 Japanese is catching up and is in fact shedding light on language production processes that could not have been uncovered if the languages studied are limited to European languages. As more aspects of language production processes in Japanese and their implications for sentence production models are clarified, SLA researchers' tasks will become more viable.

We have seen that many recent studies demonstrate that bilingual speakers mobilize linguistic resources of both languages that they possess. If SLA researchers assume the processing stages proposed by Levelt's (1989) model, then the model that de Bot adapted needs to be updated to allow vigorous interaction of two formulators. Alternatively, if Hartsuiker et al.'s (2004) model is adopted, then we assume shared combinatory nodes for structures similar in bilingual speakers' two languages and separate combinatory nodes for the structures that are dissimilar.

So far this line of research has almost exclusively been conducted on bilinguals whose two languages are English or other European languages. To my knowledge, Shin and Christianson's (2008) study, investigating English-Korean bilinguals, is the only exception. Conducting cross-linguistic priming experiments on bilingual Japanese speakers will undoubtedly help to elucidate the nature of syntactic similarities required for bilingual speakers to have shared syntax in their processing system. For example, Japanese-Korean bilinguals may have shared syntax for datives in the two languages (e.g., Japanese NP[Beneficiary]-DAT NP[Theme]-ACC may prime Korean PP [Beneficiary]-NP[Theme] rather than NP[Beneficiary]-ACC NP[Theme]-ACC) despite the differential linguistic analyses of the dative structures in two languages if the bilinguals' perceptions of similarity, rather than the linguistic analyses matter for the learners to collapse the two languages' combinatory nodes. The absence or presence of cross-linguistic priming between the different alternatives of dative structures among Japanese-Korean bilinguals may also clarify the respective roles of the order of thematic roles, syntax, and superficial resemblance of structures by comparing the effect of Korean sentences (14)–(16) above as primes for Japanese sentence production.

A model also needs to allow the Action Generator (where gesture is generated) to have access to both L1 and L2 formulators (as well as L1 and L2 lexicon). There is also some indication that the language being spoken as well as the other language being activated influence the speaker's thinking (i.e., conceptual representations of the motion events being depicted).

SLA research on L2 Japanese language production from psycholinguistic perspectives is still in its infancy, but there are now studies and models that we can consider and build on. Though it is certainly a challenging area to endeavor, at the same time it is also undoubtedly promising.

## Acknowledgments

I would like to express my sincere gratitude to the editor of this volume, Mineharu Nakayama, for not only inviting me to contribute to this volume, but also for his continuous support and helpful comments.

## References

Allen, Shanley, Asli Özyürek, Sotaro Kita, Amanda Brown, Reyham Furman, Tomoko Ishizuka and Mihoko Fujii. 2007. Language-specific and universal influences in children's syntactic packaging of manner and path: A comparison of English, Japanese, and Turkish. *Cognition* 102(1). 16–48.

Berg, Thomas. 1987. The case against accommodation: Evidence from German speech error data. *Journal of Memory and Language* 26(3). 277–299.

Bernolet, Sarah, Robert J. Hartsuiker and Martin J. Pickering. 2007. Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(5). 931–949.

Bernolet, Sarah, Robert J. Hartsuiker and Martin J. Pickering. 2009. Persistence of emphasis in language production: A cross-linguistic approach. *Cognition* 112(2). 300–317.

Bernolet, Sarah, Robert J. Hartsuiker and Martin J. Pickering. 2013. From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition* 127(3). 287–306.

Bock, Kathryn. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18(3). 355–387.

Bock, Kathryn. 1989. Closed-class immanence in sentence production. *Cognition* 31(2). 163–186.

Bock, Kathryn and Helga Loebell. 1990. Framing sentences. *Cognition* 35(1). 1–39.

Bock, Kathryn. 1996. Language production: Methods and methodologies. *Psychonomic Bulletin and Review* 3(4). 395–421.

Branigan, Holly P. and Eleonora Feleki. 1999. Conceptual accessibility and serial order in Greek speech production. In Martin Hahn and Scott C. Stoness (eds.), *Proceedings of the 21st Cognitive Science Society Conference*, 91–101. Hillsdale, NJ: Lawrence Erlbaum.

Brown, Amanda and Marianne Gullberg. 2008. Bidirectional crosslinguistic influence in L1–L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition* 30(2). 225–251.

Brown, Amanda and Marianne Gullberg. 2011. Bidirectional cross-linguistic influence in event Conceptualization? Expressions of path among Japanese learners of English. *Bilingualism: Language and Cognition* 14(1). 79–94.

Brown, Amanda and Marianne Gullberg. 2012. L1–L2 convergence in clausal packaging in Japanese and English. *Bilingualism: Language and Cognition*. DOI: http://dx.doi.org/10.1017/S1366728912000491, published online November 2012. Accessed on 22 January 2013.

Cadierno, Teresa. 2008. Learning to talk about motion in a foreign language. In Peter Robinson and Nick C. Ellis (eds.), *Handbook of cognitive linguistics and second language acquisition*, 239–275. London: Routledge.

Chang, Franklin. 2009. Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language* 61(3). 374–397.

Chang, Franklin. 2015. The role of learning in theories of English and Japanese sentence processing. In Mineharu Nakayama (ed.), *The handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Chang, Franklin, Kathryn Bock and Adele. E. Goldberg. 2003. Can thematic roles leave traces of their places? *Cognition* 90(1). 29–49.

Choi, Soonja and Melissa Bowerman. 1991. Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* 41(1). 83–121.

Choi, Soojung and James P. Lantolf. 2008. Representation and embodiment of meaning in L2 communication: Motion events in the speech and gesture of advanced L2 Korean and L2 English speakers. *Studies in Second Language Acquisition* 30. 191–224.

Cook, Vivian J. 1992. Evidence for multicompetence. *Language Learning* 42(4). 557–591.

Costa, Albert. 2005. Lexical access in bilingual production. In Judith F. Kroll and Annette M. B. de Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches*, 308–325. New York: Oxford University Press.

Costa, Albert, F.-Xavier Alario and Núria Sebastián-Gallés. 2009. Cross-linguistic research on language production. In M. Gareth Gaskell (ed.), *Oxford handbook of psycholinguistics*: 531–546. New York: Oxford University Press.

de Bot, Kees. 1992. A bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics* 13(1). 1–24.

Dell, Gary S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93(3). 283–321.

Eberhard, Kathleen M. 1999. The accessibility of conceptual number to the processes of subject–verb agreement in English. *Journal of Memory and Language* 41(4). 560–578.

Ferreira, Victor S., and Hiromi Yoshita. 2003. Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research* 32(6). 669–692.

Ferreira, Victor S., and Robert L. Slevc. 2007. Grammatical encoding. In M. Gareth Gaskell (ed.), *Oxford handbook of psycholinguistics*, 453–469. New York: Oxford University Press.

Garrett, Merrill F. 1975. The analysis of sentence production. In Gordon H. Bower (ed.), *The psychology of learning and motivation: Advances in research and theory*, 133–177. New York: Academic Press.

Garrett, Merrill F. 1988. Processes in language production. In Frederick J. Newmeyer (ed.), *Language: Psychological and biological aspects*, 69–96. (Linguistics: The Cambridge Survey.) New York: Cambridge University Press.

Goldberg, Adele. E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Hamano, Shoko. 1997. *Sound-symbolic system of Japanese*. Tokyo: Kurosio Publishers.

Hartsuiker, Robert, J. 2013. Bilingual strategies from the perspective of a processing model. *Bilingualism: Language and Cognition* 16(4). 737–739.

Hartsuiker, Robert. J., Herman H. J. Kolk and Wendy J. Huinck. 1999. Agrammatic production of subject–verb agreement: The effect of conceptual number. *Brain and Language* 69(2). 119–160.

Hartsuiker, Robert J. and Martin J. Pickering. 2008. Language integration in bilingual sentence production. *Acta Psychologica* 128(3). 479–489.

Hartsuiker, Robert J., Martin J. Pickering and Eline Veltkamp. 2004. Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science* 15(6). 409–414.

Hernández, Arturo. E., Eva M. Fernández and Noemí Aznar-Besé. 2007. Bilingual sentence processing. In M. Gareth Gaskell (ed.), *Oxford handbook of psycholinguistics*, 371–384. New York: Oxford University Press.

Hopper, Paul J. and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56 (2): 251–299.

Inoue, Kazuko. 1998. On the Japanese particle *o*. *Trends in Linguistics Studies and Monographs* 116. 449–472.

Iwasaki, Noriko. 2003. L2 acquisition of Japanese: Knowledge and use of case particles in SOV and OSV sentences. In Simin Karimi (ed.), *Word order and scrambling*, 273–300. Oxford, UK: Blackwell Publishing.

Iwasaki, Noriko. 2006a. Transitivity in Japanese sentence production: Speech errors of the dative *NI* and the accusative *O*. *Journal of Japanese Linguistics* 22. 43–57.

Iwasaki, Noriko. 2006b. Processes in L2 Japanese sentence production. In Mineharu Nakayama, Reiko Mazuka and Yasuhiro Shirai (eds.), *The handbook of East Asian psycholinguistics: Volume 2, Japanese*, 158–164. Cambridge, UK: Cambridge University Press.

Iwasaki, Noriko. 2007. Case particle errors in Japanese: Is the nominative *ga* a default case marker in sentence production? In Carson T. Schütze and Victor S. Ferreira (eds.), *The state of the art in speech error research: Proceedings of the LSA Institute Workshop* (MIT Working Papers in Linguistics 53), 205–219. Cambridge, MA: MIT Department of Linguistics.

Iwasaki, Noriko. 2011. Incremental sentence production: observations from elicited speech errors in Japanese. In Hiroko Yamashita, Yuki Hirose and Jerome L. Packard (eds.), *Processing and producing head-final structures* (Studies in theoretical psycholinguistics 38), 131–151. Dordrecht: Springer.

Iwasaki, Noriko. 2013. Describing motion in L2 Japanese: Rolling, flying up and crashing. Paper presented at the Grammar of Mimetics Workshop, SOAS, University of London, London. 10–11, May.

Iwasaki, Noriko, Gabriella Vigliocco and Merrill F. Garrett. 1998. Adjectives and adjectival nouns in Japanese: Psychological processes in sentence production. In David J. Silva (ed.), *Japanese/Korean Linguistics*, Vol. 8, 93–106. Stanford, CA: Center for the Study of Language and Information.

Iwasaki, Noriko, David. P. Vinson, Gabriella Vigliocco, Masumi Watanabe and Joanne Arciuli. 2008. Naming action in Japanese: Effects of semantic similarity and grammatical class. *Language and Cognitive Processes* 23(6). 889–930.

Jaeger, T. Florian and Elisabeth J. Norcliffe. 2009. The cross-linguistic study of sentence production. *Language and Linguistics Compass* 3(4). 866–887.

Kakehi, Hisao, Ikuhiro Tamori and Lawrence Schourup. 1996. *Dictionary of iconic expressions in Japanese K-Z* (Trends in linguistic documentation 12). Berlin: Mouton de Gruyter.

Kempen, G. and E. Hoenkamp. 1987. An Incremental Procedural Grammar for sentence formulation. *Cognitive Science* 11(2). 201–258.

Kita, Sotaro. 1997. Two-dimensional semantic analysis of Japanese mimetics. *Linguistics* 35(2). 379–416.

Kita, Sotaro. 2001. Semantic schism and interpretive integration in Japanese sentences with a mimetic: a reply to Tsujimura. *Linguistics* 39(2). 419–436.

Kita, Sotaro and Asli Özyürek. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48(1). 16–32.

La Heij, Wido. 2005. Selection processes in monolingual and bilingual lexical access. In Judith F. Kroll and Annette M. B. de Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches*, 289–307. New York: Oxford University Press.

Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, Willem J. M., Ardi Roelofs and Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22. 1–75.

Loebell, Helga and Kathryn Bock. 2003. Structural priming across languages. *Linguistics* 41. 791–824.

Mayer, Mercer. 1969. *Frog, Where are you?* New York: Dial Press.

McDonald, Janet L., Kathryn Bock and Michael H. Kelly. 1993. Word and world order: Semantic, phonological and metrical determinants of serial position. *Cognitive Psychology* 25. 188–230.

Meijer, Paul J. A. and Jean E. Fox Tree. 2003. Building syntactic structures in speaking: A bilingual exploration. *Experimental Psychology* 50(3). 184.

Melinger, Alissa, Thomas Pechmann and Sandra Pappert. 2009. Case in language production. In Andrej Malchukov and Andrew Spencer (eds.), *The Oxford handbook of CASE*, 385–401. New York: Oxford University Press.

Meyer, Antje S. and Eva Belke. 2007. Word form retrieval in language production. In M. Gareth Gaskell (ed.), *The Oxford handbook of psycholinguistics*, 471–487. New York: Oxford University Press.

Mori, Yoshiko and Junko Mori. 2011. Review of recent research (2000–2010) on learning and instruction with specific reference to L2 Japanese. *Language Teaching* 44(4). 447–484.

Nicol, Janet L., Matthew Teller and Delia Greth. 2001. Production of verb agreement in monolingual, bilingual and second-language speakers. In Janet L. Nicol (ed.), *One mind, two languages: Bilingual language processing*, 117–133. Malden, MA: Blackwell Publishers.

Nuckolls, Janis B. 1999. The case for sound symbolism. *Annual Review of Anthropology* 28. 225–252.

O'Grady, William, Yoshie Yamashita and Sun-young Lee. 2005. A note on canonical word order. *Applied Linguistics* 26 (3): 453–458.

Perniss, Pamela, Robin. L. Thompson and Gabriella Vigliocco. 2010. Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology* 1. 1–15.

Pickering, Martin J. and Holly P. Branigan. 1998. The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language* 39(4). 633–651.

Potter, Mary C. and Linda Lombardi. 1998. Syntactic priming in immediate recall of sentences. *Journal of Memory and Language* 38(3). 265–282.

Poulisse, Nanda. 1997. Language production in bilinguals. In Anette M. B. de Groot and Judith F. Kroll (eds.), *Tutorials in bilingualism*, 201–224. Mahwah, NJ: Lawrence Erlbaum.

Poulisse, Nanda. 1999. *Slips of the tongue: Speech errors in first and second language production* (Studies in bilingualism Vol. 20). Amsterdam: John Benjamins Publishing Company.

Poulisse, Nanda and Theo Bongaerts. 1994. First language use in second language production. *Applied Linguistics* 15(1). 36–57.

Sadakane, Kumi and Masatoshi Koizumi. 1995. On the nature of the 'dative' particle *ni* in Japanese. *Linguistics* 33(1). 5–34.

Schiller, Niels O., Victor S. Ferreira and F.-Xavier Alario. 2007. Words, pauses, and Gestures: New directions in language production research. *Language and Cognitive Processes* 22(8). 1145–1150.

Schriefers, H., Antje S. Meyer and Willem J. M. Levelt. 1990. Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language* 29. 86–102.

Shin, Jeong-Ah and Kiel Christianson. 2009. Syntactic processing in Korean–English bilingual production: Evidence from cross-linguistic structural priming. *Cognition* 112(1). 175–180.

Slobin, Dan I. 1996. From 'thought and language' to 'thinking for speaking'. In John J. Gumperz and Stephen C. Levinson (eds.), *Rethinking linguistic relativity*, 70–96. Cambridge, UK: Cambridge University Press.

Stemberger, Joseph Paul. 1982. The nature of segments in the lexicon: Evidence from speech errors. *Lingua* 56(3). 235–259.

Tai, James H-Y. 1985. Temporal sequence and Chinese word order. In John Haiman (ed.), *Iconicity in syntax, 49–72*. Amsterdam: John Benjamins.

Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. *Language Typology and Syntactic Description* 3. 57–149.

Talmy, Leonard. 2000. *Toward a cognitive Semantics: Typology and process in concept structuring Vol. 2*. Cambridge, MA: MIT Press.

Tanaka, Mikihiro N., Holly. P. Branigan, Janet F. McLean and Martin J. Pickering. 2011. Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language* 65(3). 318–330.

Terao, Yasushi. 1995. Bunsanshutsu-katei ni okeru tōgobumon-kenkyū no tenbō: Joshi no hatsuwa dēta o shiryō to shite [Outlook of research on syntactic aspects of sentence production: Evidence from case particle production data]. *Bulletin of Tokoha Gakuen Junior College* 26. 245–255.

Ullman, Michael T. 2001. The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition* 4. 105–122.

Vigliocco, Gabriella, Tiziana Antonini and Merrill F. Garrett. 1997. Grammatical gender is on the tip of Italian tongues. *Psychological Science* 8(4). 314–317.

Vigliocco, Gabriella, Brian Butterworth and Merrill F. Garrett. 1996. Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition* 61(3). 261–298.

Vigliocco, Gabriella., Brian Butterworth and C. Semenza. 1995. Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language* 34(2). 186–215.

Vigliocco, Gabriella and Julie Franck. 1999. When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language* 40(4). 455–478.

Vigliocco, Gabriella and Robert J. Hartsuiker. 2002. The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin* 128(3). 442–472.

Vigliocco, Gabriella and Sotaro Kita. 2006. Language-specific properties of the lexicon: Implications for learning and processing. *Language and Cognitive Processes* 21(7–8). 790–816.

Vigliocco, Gabriella, David P. Vinson, William Lewis and Merrill F. Garrett. 2004. Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology* 48(4). 422–488.

Vigliocco, Gabriella, David P. Vinson, Randi C. Martin and Merrill F. Garrett. 1999. Is 'count' and 'mass' information available when the noun is not? An investigation of tip of the tongue states and anomia. *Journal of Memory and Language* 40(4). 534–558.

Vigliocco, Gabriella, David P. Vinson and Simona Siri. 2005. Semantic similarity and grammatical class in naming actions. *Cognition* 94(3). B91–B100.

Vinson, David P., Robin L. Thompson, Robert Skinner and Gabriella Vigliocco Submitted. A faster path to meaning? Iconicity affects comprehension and production in British Sign Language. Unpublished manuscript.

Whorf, Benjamin. L. 2012 [1940]. Linguistics as an exact science. In John B. Carroll, Stephen C. Levinson and Penny Lee (eds.), *Language, thought and reality: Selected writing of Benjamin Lee Whorf. 2nd Edn, 281–298 reprinted from Technol. Rev. 43: 61–63)* Cambridge, MA: MIT Press.

Yamashita, Hiroko and Franklin Chang. 2006. Sentence production in Japanese. In Mineharu Nakayama, Reiko Mazuka and Yasuhiro Shirai (eds.), *The handbook of East Asian psycholinguistics: Volume 2, Japanese*, 291–297. New York: Cambridge University Press.

Yamashita, Hiroko, Franklin Chang and Yuki Hirose. 2003. Language-dependent aspects of structural priming. Poster presented at the Annual Conference on Architectures and Mechanisms for Language Processing, Glasgow, UK. 26 August.

Yoshioka, Keiko and Eric Kellerman. 2006. Gestural introduction of Ground reference in L2 narrative discourse. *IRAL* 44. 173–195.

Katsuo Tamaoka

# 18 Processing of the Japanese language by native Chinese speakers

## 1 Introduction

A great number of native Chinese speakers have been learning Japanese as a foreign language. According to the Japan Foundation (Kokusai Kōryū Kikin) (2011), the numbers of Japanese learners in 2009 were: 2,362 at elementary schools, 59,526 at secondary schools, and 529,508 at higher education institutions in mainland China, and 2,440 at elementary schools, 77,139 at secondary schools, and 119,898 at higher education institutions in Taiwan. Out of 128,161 foreign nationals who studied Japanese in Japan in 2011, the largest population enrolled in Japan's 1,832 higher education institutions were Chinese speakers (63,249 from mainland China and 4,134 from Taiwan) according to the Agency for Cultural Affairs in Japan (Bunka-chō) (2011). Approximately half (52.58%) of the total learners studying Japanese in Japan were estimated to be native Chinese speakers. As the number of learners increases, various issues have been identified in their processing of Japanese. However, many studies regarding these issues have been published in journals in Japan, the majority of which are in Japanese. Given the nature of this handbook, I will introduce a variety of Japanese publications to English-speaking audiences, including the latest studies on lexical pitch accent, lexical access, and sentence processing by Chinese speaking learners of Japanese. While doing so, I will clarify the ultimate goals and issues of current second language processing research. The organization of this chapter is as follows: Section 2 discusses studies on lexical pitch accents and Section 3 provides a summary of studies on processing *kanji* compounds. Morpho-syntactic processing will be discussed in Section 4 and finally, the summary of this chapter and future challenges will be provided in Section 5.

## 2 Activation of lexical pitch accents

Tones in Mandarin Chinese (considered standard Chinese) spoken in Beijing are assigned to each syllable corresponding to a Chinese character. In contrast, Japanese pitch accents are fixed to each one of the moras in a word. Japanese pitch accent is linguistically said to be an attribute of lexical items (e.g., Sugito 1982, 1989; Taylor 2011a, 2011b). In this sense, Japanese pitch accents are assumed to be lexically stored with phonological representations at the word level, and are then possibly activated together with the pronunciation of a word (e.g., Cutler and Otake 1999;

Sekiguchi 2006). Then, if this is true for native Japanese speakers, the question arises whether native Chinese speakers learning Japanese activate pitch accents when processing Japanese lexical items.

## 2.1 Perception of Japanese pitch accent by native Japanese speakers

The number of possible pitch accents in Japanese is the number of moras plus one (i.e., N + 1) (Sugito 1982). For instance, any word constructed of 3 moras has four different pitch accent patterns, and likewise, words of 4 moras have five. Regardless of the number of pitch accents possible, all words in the Tokyo-standard Japanese (i.e., *hyōjun-go* meaning 'the standard language') are classified into the following four patterns (Saito 2006; Vance 2008).

As shown in Figure 1, the first pattern is called *atama-taka-gata*, an example being the 4-mora-word HLLL+L (H here refers to high pitch and L refers to low pitch) *raigetu-ga* NP('next month')-NOM with the nominative case particle *–ga*. In this pattern, the first mora has a high pitch which drops on the second mora, and then levels out on the rest of moras. In fact, *atama* means 'initial', *taka* 'high', and *gata* 'pattern', so that the compound word literally means 'the initial-mora high pitch pattern'. The second pattern is called *naka-taka-gata*, meaning 'the middle-mora high pitch pattern'. In this pattern, the pitch rises from low to high, drops on the

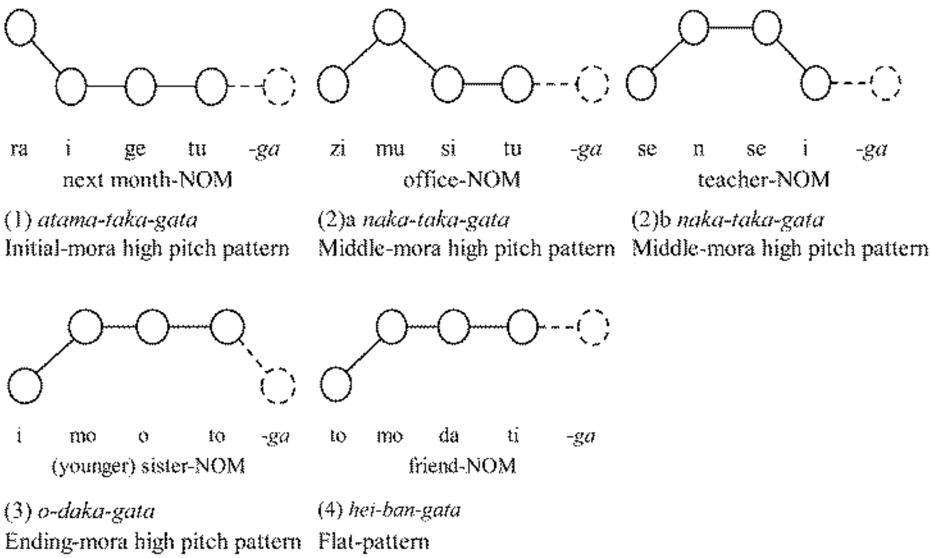



**Figure 1:** Four Japanese pitch patterns exemplified by 4-mora words with nominative case particle *-ga*

following mora, and levels out on the rest of moras. A 4-mora-word example of this pattern includes two different patterns as in LHLL+L *zimusitu-ga* NP('office')-NOM and LHHL+L *sensei-ga* NP('teacher')-NOM, depicted in Figure 1, since there are two middle moras in 4-mora words. The third pattern is called *o-daka-gata*, meaning 'the ending-mora high pitch pattern'. In this pattern, pitch rises from low on the first mora to high on the second mora, then levels out on the rest of the moras. The fourth pattern, called *hei-ban-gata* meaning 'flat-pattern', has the same pitch pattern as the third in isolated words. The third and fourth patterns can only be distinguished by means of the pitch of the following particle (Vance 2008). By adding the nominative case particle *-ga* to a noun, the third pattern of *imooto-ga*, NP('sister')-NOM is pronounced as LHHH+L whereas the fourth or flat pattern of a CV+CV+CV+CV-patterned word (where C and V refer to a consonant and a vowel, respectively), *tomodati-ga*, NP('friend')-NOM is pronounced as LHHH+H. In this manner, the pitch of the following particle changes the accent pattern.

There are some non-accented dialects sprinkled throughout Japan, in prefectures such as Miyagi, Yamagata and Fukushima. Otake and Cutler (1999) reported that Japanese pitch accents were used to distinguish the appropriate lexical items by not only speakers of the Tokyo standard dialect but also by non-accented dialect speakers. According to this finding, native Japanese speakers can fundamentally perceive pitch accent regardless of whether they are from an accented or non-accented dialect region. Yet, Otake (2002) further conducted a comparative experiment, showing that accented dialect speakers were more sensitive to pitch accents than non-accented dialect speakers. Although both speakers of accented and non-accented dialects would likely be sensitive to the Tokyo standard accent, speakers of accented dialects perceive pitch patterns more accurately than those from non-accented dialects. Overall, as indicated by the findings from native speakers until now, Japanese pitch accents seem to be activated during word processing.

Nevertheless, regional differences in pitch accent are abundant (e.g., Hattori 1951; Hirayama 1957, 1968; Kindaichi 1974; Kubozono and Ota 1998; Sugito 1982, 2006, 2012). Accents in the Osaka region often show pitch reversal compared to the Tokyo-standard accents, such as 4-mora CV+CV+φV+CV-structured (φ refers to an empty consonant) *boosi-ga* 'a hat' NP('hat')-NOM, which is LHHH in Tokyo versus HLLL in Osaka. A study on perception of the Tokyo-standard accent by speakers of different dialect backgrounds was conducted by Ayusawa (1998). This study used the pitch accent test developed by Nishinuma (1994), which has 72 items consisting of 3 to 5-mora words such as *megane* 'glasses', *katakana* 'katakana script', and *otokonoko* 'boy'. With no statistical analyses conducted on these data, Ayusawa (1998) commented that a majority of native Japanese speakers are likely to correctly perceive the Tokyo-standard accent (Ayusawa 1998: 70–71). However, this inference is misleading. Even by merely glancing at the figures of the graphs shown in Ayusawa (1998: 72), it is quite obvious that participants in the Tokyo-standard areas show

significantly higher accuracies than participants in other dialect areas on various types of words.

Furthermore, we can gather the following conclusions from Ayusawa (1998) by simply looking at the accuracy rates. Thirty participants from Ibaraki and Fukushima prefectures showed a lesser degree of accuracy on perception for the first (mora) accented items for 3-mora words, the first and the second (mora) accented words for 4-mora words, and the second and the third (mora) accented words for 5-mora accented words in comparison to 30 participants from Tokyo. Likewise, 30 participants from the Osaka and Kobe areas showed an even less accurate trend for perception of first mora accented items for 3-mora words, the first and second mora accented words for 4-mora words, and the first, second, and third mora accented words for 5-mora accented words. In contrast, it is interesting that participants from the Chugoku area showed higher accuracy across all conditions, similar to those from Tokyo. It is still an unanswered question as to whether Japanese native speakers obligatorily activate pitch accents along with lexical items due to dialectic variations.

## 2.2 Influence of Japanese language proficiency on pitch accent acquisition

Even though pitch accents show diverse differences across the regions of Japan, Tokyo-standard accents are taught intensively at a majority of Chinese universities. Widely-used Japanese textbooks for native Chinese speakers at universities in China (e.g., Hong 2010; Pan 2011; Zhang 2011; Zhao 2012; Zhou and Chen 2009, 2010, 2011a, 2011b) describe the position of a word's pitch accent when introducing Japanese vocabulary. For example, in the three-mora CV+CV+φV-structured adjective *warui* meaning 'bad', a high pitch accent is indicated as being placed on the second mora *ru*, denoting a LHL pattern. Since Japanese accents are thoroughly instructed when Chinese students learn Japanese words, Chinese learners of Japanese quite possibly memorize pitch accents as they learn new words, activating these accents when processing Japanese lexical items.

Lee, Murashima and Shirai (2006) conducted a longitudinal study on three Chinese learners of Japanese, Jane, Mary, and Ann. These three native Cantonese speakers were born and raised in Hong Kong, and were all 18 years old at the start of the research. They tested these learners' production of pitch accents at three times, December 1999, February 2001, and February 2002. They showed that these three Chinese learners of Japanese did not display any changes in production accuracies for Japanese pitch accent over three testing periods as in Table 1.

**Table 1:** Three Chinese Speaking Learners of Japanese

| | December 1999 | February 2001 | February 2002 |
|---|---|---|---|
| Jane | 67.6% | 54.3% | 70.6% |
| Mary | 70.6% | 62.9% | 74.3% |
| Ann | 74.3% | 77.1% | 70.6% |

It was also reported that the three Chinese learners of Japanese varied in their overall Japanese language proficiency at the conclusion of two years of study. Nevertheless, there were no differences in production accuracy of Japanese pitch accents among them. Thus, they concluded that these learners showed no improvement in the production of Japanese word accents during the two years. If this finding is taken as indicated, Chinese learners of Japanese would not activate standard pitch accents when processing Japanese words.

Lee, Murashima and Shirai (2006), however, used only three participants for investigating pitch accent production. Although it is not a production study, Pan (2003) conducted a study on perception of Japanese pitch accent with a larger group of native Chinese speakers. This study measured accuracy of accent perception on two-mora words. The task was conducted with three groups of native Chinese speakers studying at a university in Taiwan: 36 Japanese learners majoring in Japanese language, 30 not majoring in Japanese, and 21 native Chinese speakers with no Japanese learning experience. Accuracies of two-mora words clearly differed among the three groups, as in Table 2.

**Table 2:** Accuracies of two-mora words by group

| | Two mora words | Flat | Early high pitch | Late high pitch |
|---|---|---|---|---|
| Japanese majors | 95.36% | 98.61% | 95.84% | 91.32% |
| Non-Japanese majors | 74.86% | 76.25% | 80.00% | 68.33% |
| No Japanese learning | 48.22% | 42.86% | 60.12% | 41.67% |

Pan also reported accuracies based on the three different accent patterns, namely, the flat (*heiban*), the early high pitch (*atama-taka-gata*), and the later high pitch (*o-daka-gata*) patterns for two mora words, as in Table 2 above. The students specializing in Japanese showed very high perception accuracies whereas the non-majors showed a significant difference in ascending order from the highest in the early high pitch pattern, the flat pattern, and the later high pitch pattern. Since Japanese majors were expected to have a higher Japanese proficiency than those who were non-majors, Pan demonstrated that Japanese language proficiency contributes to higher accuracy on pitch accent perception, unlike Lee et al.'s (2006) case study. Chinese speaking Japanese majors are likely to efficiently acquire the ability to accurately perceive the Tokyo standard pitch accents.

If Japanese pitch accent is an attribute of lexical items (e.g., Sugito 1982, 1989; Taylor 2011a, 2011b), and if accent is activated during word processing (e.g., Cutler and Otake 1999; Otake 2002; Otake and Cutler 1999), accent should be stored as lexical knowledge in the mental lexicon and utilized for word processing. As we discussed above, Chinese speaking learners of Japanese in Lee et al. (2006) did not show any improvement in pitch accent production over two years of Japanese study, and yet those in Pan's (2003) study demonstrated a notable difference in perception accuracy between Japanese majors and non-Japanese majors. The differences in these studies may come from the difference between production and perception studies. To clarify the findings, it is necessary to conduct more production and perception studies on Japanese pitch accent with participants whose Japanese language proficiency levels, especially, lexical knowledge, are controlled. With the appropriate control, one can truly observe whether accuracy of pitch accent improves as Japanese proficiency increases.

## 2.3 Dialectal variation in Japan influencing pitch accent acquisition

As noted earlier, pitch accent patterns vary regionally across Japan (e.g., Hattori 1951; Hirayama 1957, 1968; Kindaichi 1974; Kubozono and Ota 1998; Sugito 1982, 2006, 2012). However, we do not know exactly how dialect differences affect accent production and perception by Chinese learners of Japanese. In an attempt to address this question, Yang (2011) conducted an interesting study on native Chinese speakers who had been studying Japanese in the Kansai region, in which accent differs greatly from the Tokyo-standard pattern. Yang (2011) compared 30 native Chinese speakers studying Japanese in Taiwan with 30 native Chinese speakers learning Japanese in the Kansai region. The study reported that Chinese learners of Japanese studying in Taiwan performed significantly better in both production and perception of the Tokyo-standard pitch accent than those studying in Kansai. This difference in accent performance may be created by a dialect accent specific to the Kansai region. However, before reaching the conclusion of dialect influence, two factors must be pointed out in Yang's study, which possibly resulted in lower accuracy. First, the Japanese proficiency of the native Chinese speakers studying in Taiwan and in the Kansai region should have been controlled because Pan (2003) showed an effect of Japanese language proficiency on perception ability. Second, native Chinese speakers studying in areas where the Tokyo-standard accent is spoken should have been contrasted with those in the Kansai area in order to directly examine a potential disadvantage of learners studying in non-standard dialect areas.

Japanese regional accents do display great diversity. Yang showed a lesser degree of accuracy among Chinese speakers studying in the Kansai region in comparison to those studying in Taiwan. If native Chinese speakers were taught

Japanese in the Tokyo-standard accent at a university or in a Japanese language school in Taiwan, they may be able to gradually memorize a single type of pitch accent. In contrast, Chinese speakers studying in the Kansai region have to face conflicting input from the environment, of which accent patterns differ from the Tokyo-standard accent. In all likelihood, these learners study new words with the Tokyo-standard accent within their Japanese classrooms. However, once they set foot outside of the classroom, they are immersed in a different accent environment. Chinese-speaking learners have to constantly face conflicting accentual input in their daily lives. Thus, it is hypothesized that a certain dialect environment whose accent greatly differs from the Tokyo-standard type would interfere with the acquisition of the Tokyo-standard accent. How dialect accents interfere with the Tokyo-standard accent is an important pedagogical issue to investigate. An ideal way to investigate the dialect interference is to have two groups of Chinese-speaking learners of Japanese sampled from the Kansai and Kanto areas and matched by Japanese language proficiency, then tested for accuracy of production and perception in the Tokyo-standard pitch accent. Moreover, since pitch accent may not be a very reliable cue for lexical access during spoken word recognition, the usefulness of pitch accents could be measured by the magnitude of contribution to listening comprehension.

## 2.4 Dialect diversity of Chinese influencing Japanese pitch accent

Dialect diversity also exists in Chinese tone accents. Tone accents in the Beijing dialect, which is considered standard Chinese (i.e., Mandarin Chinese), are put on each syllable, whereas tones in the Shanghai dialect are realized at the word level (Hayata 1999; Xu and Tang 1988; You 2004). Iwata (2001) suggests that tones in the Shanghai dialect resemble Japanese pitch accents in that both accents are realized at the word level. If this is true, Chinese learners of Japanese from the Shanghai dialect should show an advantage in acquisition of Japanese pitch accents over those of the Beijing dialect.

The influence of different Chinese dialects on the acquisition of Japanese pitch accent was investigated by Liu (2010), who tested 18 Beijing dialect participants and 21 Shanghai dialect participants. Liu conducted a production study of verb-plus-verb-structured (V+V) compound verbs (e.g., *tumi+ageru* ‘pile’ and *tori+kaesu* ‘take back’). Japanese has an abundance of compound verbs produced by combining two native Japanese (*yamato kotoba*) verbs. When two verbs are combined, a compound verb changes the position of its pitch accent. For example, the verb *tori* ‘take’ has a HL accent (*atama-taka-gata*). Another verb *kaesu* ‘return’ has an HLL pattern, which is also categorized as *atama-taka-gata*. When these two verbs are combined, the result is the compound verb *torikaesu* ‘take back’. In the compounding process, the accent of the first verb *toru* is altered to the flat accent of HH,

becoming a LHHLL accent, or the flat accent LHHHH (Liu 2010: 17). Due to the complexity of this accent variation, Japanese compound verbs are expected to be difficult to acquire for native Chinese speakers. Liu (2010), therefore, hypothesized that native Chinese speakers of the Shanghai dialect would perform better in perceiving and producing the correct pitch accents of compound verbs than those of the Beijing dialect.

The results (see Liu 2010: 18, Table 3) were rather intricate. The speakers of the Shanghai dialect produced the pitch accents of compound verbs more accurately than those of the Beijing dialect when the compound verbs were accented on the penultimate mora (or denoted as -2 accent) or the flat accent (or 0 accent). However, with compound verbs accented on the third mora (or -3 accent) or the flat accent (or 0 accent), the result was reversed in such a way that speakers of the Beijing dialect performed better than those of the Shanghai dialect. Therefore, it cannot be simply assumed that Chinese learners of Japanese whose accents are realized at the word level (i.e., the Shanghai dialect) perform better at producing pitch accents than those with tone accents at the syllable level (i.e., the Beijing dialect).

Before making any further comments, three basic methodological problems in Liu's study (2010) should be pointed out. First, the Japanese ability of native Chinese speakers of the Beijing and the Shanghai dialects was controlled as having learned Japanese for two years. As commonly observed, two years of learning does not guarantee equal levels of attainment of Japanese language ability. A preferred alternative would be to conduct a Japanese vocabulary test to balance the two groups based on lexical knowledge. Second, Liu (2010) asked five native Japanese speakers of the Tokyo dialect to evaluate the accuracy of pitch accents produced by Chinese learners of Japanese. However, there is no report of consistency and reliability of these evaluators' judgments. It is hard to imagine that all five evaluators scored the participants in the same way. Third, pitch accent accuracy was scored from 1 (disagree), 2 (slightly disagree), 3 (slightly agree) and 4 (agree). Liu (2010) assigned 'correct' for 2–4 scores and assigned 'wrong' for a 1 score. This correct/incorrect judgment would have skewed ratings toward the higher possibility for a 'correct' judgment. If Liu (2010) had used a correct-or-wrong dichotomous scale for analysis, the five evaluators could have made judgments on the basis of either 'correct' or 'incorrect'.

China is diverse in its regional dialects and accents. When Chinese speakers from different dialects meet, it is frequently observed that they cannot understand one another's speech. The diversity in Chinese dialects may possibly influence acquisition of Japanese pitch accent (if we assume some kind of L1 transfer). As Liu (2010) reported, Chinese speakers of the Beijing dialect differed in production accuracy of Japanese pitch accent from those of the Shanghai dialect. Although I regard Liu's study highly for its having dealt with the unique perspective of Chinese dialects, it had some methodological issues that concern us as pointed out earlier. Therefore, a similar comparative study should be conducted in the future. It should

be noted, however, that tone accent in the Chinese language fundamentally differs from Japanese pitch accent, so that results suggesting that Chinese dialectal differences cross-linguistically influence Japanese pitch accent should be carefully interpreted.

## 2.5 Cross-linguistic studies on Japanese pitch accent

An investigation of cross-linguistic differences in the acquisition of Japanese pitch accent conducted by Ayusawa (1998) indicated a strong effect of first language on perception of Japanese pitch accent. An interesting trend in Japanese pitch accent production was reported among native Korean speakers learning Japanese (Fukuoka 2008). When Japanese words contain a voiced plosive sound in word initial position, Korean speakers are likely to put a low pitch on the initial mora. In contrary, when Japanese words contain a voiceless plosive sound in word initial position, Korean speakers are likely to put a high pitch on the initial mora. This trend in laryngeal contrast is observed in the Korean language (Kim and Duanmu 2004). Thus, this tendency could be the result of influence adapted from the learners' mother tongue of Korean (L1 transfer).

Nevertheless, Taylor (2012) points out that it is very difficult to determine whether the accent trend is caused by a learner's mother tongue. She shows examples of some trends across multiple languages based on the examination of previous studies on pitch accent (e.g., Andreev 2002; Lee, Murashima and Shirai 2006; Nakato 2001; Toda 1999; Sukegawa 1999): i) A tendency to overuse pitch accent on the penultimate mora of a word, which is observed not only among English speaking learners of Japanese but also Korean, Chinese, and Bulgarian speaking learners, and ii) a tendency to place pitch accent on heavy syllables, which is reported among Korean, Portuguese, and Bulgarian speaking learners. In a cross-language comparison study, needless to say, an important condition would be to control the levels of Japanese language proficiency, as some speakers such as Chinese and Korean speaking learners are likely to reach a high level of Japanese language proficiency within a few years, and these languages have unique linguistic differences which could aid in the investigation of native language effects on not only Japanese pitch accent but also other features in units of lexical processing such as syllable-timed vs. mora-timed languages.

## 2.6 Contribution of pitch accent in distinguishing homophones

One of the important basic functions of pitch accent is to differentiate homophonic words and identify the proper homophone in a sequence of utterances. A homophone is a word that shares the same pronunciation with another word while differing in meaning. For example, *ame* meaning 'candy' is produced with a LH pitch,

while the segmentally identical word *ame* produced with a HL pitch becomes 'rain'. To accomplish homophonic distinction, Chinese learners of Japanese must memorize the concept of the word with the proper pitch accent. In other words, they must activate the pronunciation of the word *ame* with both its pitch accents of LH for 'candy' and HL for 'rain' to identify the intended meaning.

Mathematical linguists have calculated the possibility of distinction among Japanese homophones, and have suggested that Japanese pitch accent is not necessarily crucial for accessing lexical meaning when distinguishing homophonic words. According to Shibata and Shibata (1990), 13.57% of the homophones in Japanese are distinguished by pitch accent, while in Chinese tone accent distinguishes 71.00% of homophones. Given this difference, they claimed that tone in Chinese is used for distinguishing homophones while pitch in Japanese is not. Shibata and Shibata (1990) propose only a minor role for Japanese pitch accent in homophone distinction.

Furthermore, Kitahara (2006) investigated the distribution of homophonic pairs distinguished by pitch accent (i.e., accentual oppositions) in Japanese, using the lexical database of Amano and Kondo (1999, 2000).[1] Using Amano and Kondo's (2000) frequency index, Kitahara (2006) also pointed out that homophonic minimal pairs include those which greatly differ in frequency such as /hito/ for 'human', counted 121,162 times and /hito/ as 'use of an expense', counted only twice. Therefore, once these frequency-divergent homophone pairs, which native Japanese speakers are unlikely to know, or at least will not contrast by pitch accent, were excluded, the pitch distinguishability rate of accentual oppositions will drop to less than 10 percent.

To check whether native Chinese speakers really distinguish homophones by pitch accent in Japanese, it is possible to run the following test. A sentence with the correctly-accented target word based on the Tokyo-standard accent should be presented as in (5a), where the underlined word is the target word.

(5) a. *Kodomo ni mainiti ame o katte ageteiru.*
child DAT everyday candy ACC buy give
'Everyday [I] buy candy for children.'

b. *Kodomo ni mainiti ame o katte ageteiru.*
child DAT everyday rain ACC buy give
'#Everyday [I] buy rain for children.'

In the same sentence, the LH-accented target word *ame* 'candy' is replaced, as in (5b) by changing *ame* to HL. Of course, we do not give 'rain' to children, so the

1 This database was created from the corpus of the *Asahi Newspaper* from 1985 to 1998, which contains 341,771 words for type frequency and 287,792,797 words for token frequency. Within the high familiarity range (familiarity index taken from Amano and Kondo 1999) the unaccented-accented opposition or flat pattern versus other accented patterns is more prevalent than the accent-location oppositions.

word ‘rain’ is incorrectly matched with the semantic context of the sentence. A set of homophone pairs can be used to investigate whether native Chinese speakers really activate pitch accent when accessing the concept of a lexical item.

# 3 Processing of Japanese kanji and their compound words

The writing system of the modern Japanese language consists of the *kanji* and *kana* scripts (for detail, see Hadamitzky and Spahn 1981; Kess and Miyamoto 1999; Miller 1967; Tamaoka 1991). Kanji are logographic morphological units, adapted from the script of the Chinese language. In contemporary Japanese, kanji represent not only lexical items originating from Chinese (*kango*) but also Japanese (*wago*), which were created by Japanese speakers. Kanji-compound words are extremely common in Japanese. Token frequencies of kanji-compound words encompass 41.3% of all Japanese vocabulary, as reported by Kokuritsu Kokugo Kenkyujo (1964). More dramatically, kanji compounds make up approximately 70% of the entries in a typical Japanese dictionary (Yokosawa and Umeda 1988). A kana symbol is a phonogram which fundamentally represents a single mora on a one-to-one basis. The kana script further consists of two orthographies, *hiragana* and *katakana*. The hiragana script is cursive in shape (あ for /a/) and used for grammatical morphemes as well as for some content words. The katakana script is angular in shape (ア for /a/), and usually used for writing loanwords from alphabetic languages, as well as the names of animals and plants. The hiragana and katakana scripts describe Japanese sounds on the basis of mora-to-kana correspondence. The three scripts – kanji, hiragana, and katakana – are simultaneously used in modern written Japanese texts.

A great number of Japanese *kanji* have visually similar shapes as the Chinese characters from which they were originally derived. Among a selection of 4,600 Japanese kanji-compound words, Chen (2002) counted 54.5% in mainland China and 55.1% in Taiwan that are written with the same characters and imply the same meaning as their Chinese counterparts. Additionally, 14.9% of these words in mainland China and 13.3% in Taiwan have the same characters and similar meanings, and only 4.1% in mainland China and 3.5% in Taiwan share the same characters but different meanings. In total, 73.5% of the kanji compounds used in mainland China and 71.9% in Taiwan share the same characters in both Chinese and Japanese. Moreover, Hishinuma (1983, 1984) further comments that, if the slight differences in orthographic shapes between Chinese and Japanese are ignored, it can be assumed that native Chinese speakers know 98.1% of the commonly-used Japanese kanji prior to learning the Japanese language. This great similarity of kanji morphemic units explains the commonly-observed tendency of native Chinese speakers to depend heavily on kanji meanings to understand written Japanese texts.

## 3.1 Advantage of kanji orthographic similarity in lexical processing

In studies on English as a second language (ESL), knowledge of 98% of the words in a written text is required to achieve accurate understanding of the text (Hu and Nation 2000; Nation 2001; Stahl and Nagy 2006). In Japanese, Komori, Mikuni and Kondo (2004) indicated that knowledge of 96% of the words in a written text is necessary for comprehension. This figure implies that the threshold for an appropriate level of reading comprehension would entail that less than 4% of the vocabulary in a given text be unknown. Since many Japanese words are shared with Chinese as indicated by the numbers presented above, native Chinese speakers are expected to have a great advantage in reading comprehension. Then, how much of an advantage do native Chinese speakers have in the processing of kanji-compound words? First, let's compare them with native English speakers who have no kanji knowledge. Tamaoka (1997) measures the difference in processing efficiency (i.e., speed and accuracy) of lexical decisions for Japanese kanji-compound words by 10 native Chinese and 17 native English speakers studying Japanese from two to three years at the same university in Canada under the same curriculum.[2] A great difference between the two groups was found as in Table 3.

**Table 3:** Mean response times and accuracy rates by group

| | Reading times (milliseconds) | Accuracy |
|---|---|---|
| Chinese speaking learners | 982 ms | 71.3% |
| English speaking learners | 1,808 ms | 63.7% |

Native Chinese speakers performed 826 milliseconds faster and 7.6% more accurately than native English speakers. Interestingly, the Chinese group processed two-kanji compound words with both few and many strokes equally well, whereas the English group were slower in processing compounds with many strokes than those with fewer strokes.

A script advantage for learners of Japanese with different script backgrounds was also investigated by Tamaoka (2000). He examined the effects of L1 scripts when native Chinese and English speakers phonologically processed the same Japanese words presented in three different scripts of kanji, hiragana and *romaji* (alphabetic transcription of Japanese). Fifteen native Chinese speakers and 13 native English speakers learning Japanese participated in the study; all studied Japanese for two to three years under the same curriculum at a university in Australia. The Chinese students all came from China as international students. A summary of the results are in Table 4.

---

**2** It should be noted that these Chinese university students are native Chinese speakers who came from a country where Chinese is spoken. They are sometimes referred to as 'visa students'.

**Table 4:** Mean naming times and accuracy rates by group

| | Naming times (milliseconds) | | | Accuracy | | |
|---|---|---|---|---|---|---|
| | Kanji | Hiragana | Romaji | Kanji | Hiragana | Romaji |
| Chinese group | 1,027 ms | 1,098 ms | 1,295 ms | 89.52% | 99.05% | 89.52% |
| English group | 1,635 ms | 1,009 ms | 783 ms | 53.85% | 94.51% | 95.60% |

Average naming latencies (the time from visual presentation of a word to initialization of its pronunciation) of 21 Japanese words presented in kanji (e.g., 会話, /kaiwa/ 'conversation') were faster with a higher accuracy for the Chinese group than the English group. A clear advantage for Chinese speakers was demonstrated by the difference in overall performance of 608 ms and 33.77%. In striking contrast to the case of kanji, the same words presented in hiragana (e.g., かいわ) yielded nearly identical processing performance in both language groups. Furthermore, the same words in romaji (e.g., *kaiwa*) displayed an opposite trend. The L1 script (i.e., the familiarity of the script) exhibited strong effects on phonological processing of L2 Japanese words, facilitating the processing of kanji compounds for the Chinese group and romaji for the English group. As Djojomihardjo, Koda and Moates (1994) indicated in L2 English learners, script consistency between L1 and L2 strongly facilitates the speed of L2 lexical and text processing.

Yamato and Tamaoka (2009) conducted a lexical decision task with 21 Chinese speaking learners of Japanese with higher lexical knowledge and 18 with lower lexical knowledge based on a vocabulary test (for details of the test, see Miyaoka, Tamaoka and Sakai 2011). Both proficiency groups had been learning Japanese in Japan. This study was analyzed as a 2 (participants' lexical knowledge; higher and lower lexical groups) × 2 (kanji-compound words; high- and low-frequency) design. A summary of the results are stated in table 5.

**Table 5:** Mean response times and accuracy rates of kanji-compound words by group

| | High- frequency | | Low-frequency | |
|---|---|---|---|---|
| | Response times | Accuracy | Response times | Accuracy |
| Higher lexical knowledge | 754 ms | 98.1% | 937 ms | 90.7% |
| Lower lexical knowledge | 760 ms | 97.2% | 976 ms | 78.3% |

Both groups of higher and lower Japanese lexical knowledge processed high-frequency kanji-compound words and low-frequency ones at almost the same speed. Of interest were the results of the processing accuracy measure. Although both groups with higher and lower Japanese lexical knowledge processed high-frequency kanji-compound words with high accuracy, the group with lower Japanese lexical knowledge showed lower accuracy than the group with higher lexical knowledge.

They found that the response times between the higher and lower lexical groups showed no difference whereas the higher lexical group performed more accurately on low frequency words than the lower lexical group. Regardless of Japanese word frequency and lexical knowledge, all native Chinese speakers seemed to be able to process Japanese kanji compounds quickly using their first language knowledge of Chinese characters. However, kanji compounds used in Japanese lexical items occasionally differ from their semantic usages in Chinese, which would predictably result in lower accuracy. These Japanese words tend to be low frequency words among native Chinese speakers who have not acquired the large Japanese vocabulary.

Their finding can be explained in the framework of lexical processing as follows. Native Chinese speakers quickly reach orthographic activation of a two-kanji compound word based on their (L1) character knowledge, which further activates its concept. Then, they have to determine whether this compound word really exists in the Japanese lexicon. At this stage, their Japanese knowledge of concepts begins to influence their lexical decision. If they do not have sufficient lexical knowledge of Japanese two-kanji compounds, they have no way to correctly determine the existence of the target word. Therefore, while fast speed for lexical processing was accomplished by quick activations of orthographically interconnected representations of the two languages, the difference in accuracy was created by conceptual lexical knowledge of the Chinese speakers. This can be supported by the case of native English speakers who displayed slower response times and lower accuracy for lexical decisions on two-kanji compound words (Tamaoka 1997), because the English speakers have no kanji orthographic knowledge in their first language. The English speaking learners' slow processing of Japanese words must be caused by a slower bottom-up processing which involves orthographic analysis of kanji elements, activation of each kanji, combining two kanji, and finally activating its lexical concept.

## 3.2 Advantages and disadvantages of kanji orthographic similarity in text understanding

The importance of lexical knowledge for *text understanding* is well-documented not only in English as a second language (ESL) but also in Japanese as a second/third language (JSL). Text understanding refers to the skills necessary to comprehend a text which is presented visually for reading and aurally for listening. An advantage of native Chinese speakers' knowledge of Japanese kanji in text understanding was found in tests conducted by Matsunaga (1999) at a university in southern California. She tested 12 Chinese students (one had insufficient English ability, and was excluded from the analysis) and 28 students with non-kanji backgrounds including three Spanish speakers, two Korean speakers, and one Thai speaker. A summary of the

results are stated in Table 6, where the maximum comprehension scores were 100 points.

**Table 6:** Mean comprehension scores and reading times by group and passage type

| | Narrative passages Comprehension reading times | | Descriptive passages Comprehension reading times | |
|---|---|---|---|---|
| Chinese speakers ($n$ = 11) | 89.00 | 234.00 sec | 85.47 | 129.27 sec |
| Non-kanji background ($n$ = 28) | 79.09 | 333.53 sec | 61.05 | 258.32 sec |

Students with a kanji background showed significantly higher comprehension scores and faster oral reading speed for narrative passages than students with a non-kanji background. Likewise, students with a kanji background showed significantly higher comprehension scores and faster oral reading speed for descriptive passages than those with a non-kanji background. As such, a clear tendency towards an advantage for Chinese students (i.e., kanji background) was observed among the learners at a university in an English speaking country. Matsunaga (1999), however, used English translations to check participants' understanding of the Japanese text, so that English ability must have influenced their performance.

Advantages in kanji processing by native Chinese speakers were also shown in the on-line processing of Japanese text comprehension by Yamato and Tamaoka (2013). In their study, 20 matched pairs of native Chinese and Korean speakers were selected so that they were equal in both lexical and grammar skills. This sampling method is called *pair-matched sampling*. In this method, each pair of native Korean and Chinese speakers learning Japanese at a university in their own country from two to three years was made by matching scores on two tests, a Japanese vocabulary test (maximum 48 points) and a grammar test (maximum 36 points). Participants were selected so that the average scores matched exactly between the two groups (see Table 7): The vocabulary test was exactly matched at the same average between 20 native Chinese speakers and 20 native Korean speakers. The grammar test scores also displayed nearly the same average between native Chinese speakers and native Korean speakers. This approach guarantees a direct comparison of the two different linguistic groups.

Using the fixed-window *self-paced reading* technique, the selected native Chinese and Korean groups were asked to read two texts, one with many kanji words, and one with many katakana words. In the fixed window self-paced reading, each phrase is presented to a participant one at a time in the center of a computer monitor. When a participant presses the space bar, the next phrase is displayed in the same position on the screen, and this process continues until the whole story of the text has been displayed. The time between each press of the space bar is considered to be the time required for reading each phrase. Some weaknesses of this method should be noted,

however, in that participants performing self-paced reading cannot re-read a text once they press the space bar. In addition, participants may be able to read a phrase faster than the time it takes to press the space bar.

**Table 7:** Mean test scores and reading times for kanji and katakana words by group

| | Vocabulary (SD) | Grammar (SD) | Kanji | Katakana |
|---|---|---|---|---|
| Chinese speakers (*n* = 20) | 37.90 (4.90) | 32.40 (2.93) | 1,227 ms | 2,104 ms |
| Korean speakers (*n* = 20) | 37.90 (5.60) | 32.90 (2.81) | 1,741 ms | 1,716 ms |

Due to the great similarity of Japanese kanji and Chinese characters, native Chinese speakers processed visually-presented kanji compound words in a text much faster than native Korean speakers. For example, *gyoosei too kara* 'from such areas as administration', which consisted of three kanji ('such areas as administration') and two hiragana ('from') embedded in the text, was processed with a difference of 514 ms between the two groups (see Table 7 above). In contrast, native Korean speakers processed katakana-presented alphabetic loanwords faster than native Chinese speakers. The phrase *konbiniensu sutoaa de* 'at the convenience store' written with 10 katakana ('the convenience store') and one hiragana ('at') was processed 388 milliseconds faster by Korean learners than by Chinese learners. The similarity of phonetic symbols (the symbol-to-sound conversion) between Japanese kana and Korean Hangeul scripts may have helped native Korean speakers process alphabetic loanwords quicker than native Chinese speakers. In other words, Koreans can quickly convert kana-to-sound since they frequently experience this similar conversion process in their Hangeul script. The script similarity between L1 Japanese and L2 Chinese/Korean created a diverging pattern of differences in lexical processing speed of Japanese – Chinese were superior at processing kanji compound words, while Koreans were better at alphabetic loanwords, even embedded in a text.

Sharing a majority of Japanese kanji with Chinese characters is not always beneficial (Tamaoka 1997, 2000). Due to the great resemblance of kanji, native Chinese speakers heavily rely on orthography to process two-kanji compound words in accessing their meanings. They are, in turn, likely to pay little attention to the phonological aspect of kanji compound words to understand a spoken text, as observed by the misunderstanding and dropping of information in listening comprehension (e.g., Hong 2004; Ishida 1986; Komori 2005; Yin 2002). Due to strong ties between orthography and concepts (or semantics) in kanji and their compounds, it may be the case that native Chinese speakers learning Japanese only establish weak connections from orthography to phonology.

A cross-linguistic comparison of reading and listening comprehension by native Chinese and Korean speakers learning Japanese was conducted by Komori (2005). Chinese showed a large discrepancy of 12.75% between 66.09% in reading comprehension and 53.30% in listening comprehension whereas Korean showed only a

small difference of 3.61% between 75.09% in reading comprehension and 78.70% in listening comprehension.

This study, however, contains two essential methodological problems. First, reading and listening comprehension tests were not conducted on the same Chinese and Korean groups. The reading comprehension test was conducted with 22 native Chinese speakers and 39 native Korean speakers learning Japanese at a private university in Tokyo. However, the listening comprehension test was conducted with 9 native Chinese speaking and 15 native Korean speaking participants recruited from students at the same university. These two groups were fundamentally different, so that Komori has to make an unproved assumption that two paired-groups of Koreans and Chinese are equivalent in Japanese proficiency. Second, texts used for reading comprehension differed from those for listening comprehension. The level of lexical difficulty in the texts used by the researcher was controlled according to lexical levels on the former Japanese Proficiency Test (Japan Foundation and Japan Educational Exchange and Services 2002). However, it is not ideal to use different texts to compare results of reading and listening comprehension. It would be desirable to see a similar study conducted on the same group, especially native Chinese speakers, by counterbalancing texts for reading and listening comprehension. Putting these issues aside, the study roughly depicted a cross-linguistic difference in reading and listening comprehension between native Chinese and Korean speakers learning Japanese. See also Sawasaki (2006).

## 3.3 Effects of kanji orthographic similarity between Japanese and Chinese

Chinese characters used in mainland China have undergone simplification. Soon after the foundation of the People's Republic of China on October 1, 1949, the movement to implement simplified Chinese characters got underway. A draft of the simplified Chinese character list was announced in 1955, and the first newspaper using simplified Chinese characters was published the following year, in 1956. In 1964, the Chinese government combined the simplified characters into the Total List of Character Simplification or *Jian Hua Zi Zhong Biao*. This list was reformed a few times, and the Chinese government has been collecting public comments for a modified list of simplified characters since 2009 (see details, Endo 1986). The series of simplifications resulted in some orthographic differences between Chinese characters and Japanese kanji.

Kayamoto (1995a) measured orthographic similarity or difference between Chinese characters and Japanese kanji on a scale of 0 to 4. Characters given a 0 are identical in Chinese and Japanese, such as 常 and 道. One is given to a difference of only a dot or a line of a character (歩 and 海 for Japanese, and 步 and 海 for Chinese). Two refers to a difference of a part in a character (話 and 許 for Japanese, and 话

and 许 for Chinese). Next, three indicates a large difference of a part or both sides of a character (練 and 動 for Japanese, and 练 and 动 for Chinese). Finally, four represents a complete difference of the entire character (專 and 異 for Japanese, and 专 and 异 for Chinese). A correlation between this 0-to-4 scale and subjective character-difference judgments by native Chinese speakers was reported to be very high, at 0.90 ($p$ < .001), by Kayamoto (1995a). Thus, Kayamoto's scaling seems to be reliable to use as an index for orthographic similarity between Chinese and Japanese.

Using the 0-to-4 scale, Kayamoto (1996) investigated effects of the orthographic similarity in processing of Japanese two-kanji compound words. Naming latency, which was defined as the latency from the onset of visual-presentation of a stimulus item to the offset of the first amplitude in its pronunciation, indicated that Japanese kanji compounds similar in both languages (M = 597 ms) were named faster in Chinese sounds than dissimilar ones (M = 669 ms) by native Chinese speakers leaning Japanese at the advanced level. After the naming task, these participants reported that they pronounced these Japanese kanji as if newly-simplified Chinese characters. Unlike Chinese pronunciations, when they were asked to name the same kanji compounds in Japanese, no difference was observed between those similar in both languages (M = 1,101 ms) and those that are dissimilar (M = 1,080 ms). As seen in the difference between the processing of Chinese and Japanese sounds, Japanese kanji orthographic units are strongly mapped onto Chinese sounds, even though they are not exactly identical to Chinese characters. On the contrary, these kanji are less tightly mapped onto Japanese sounds regardless of orthographic similarity.

A null effect of orthographic similarity on naming was confirmed by Kayamoto (2002). Orthographic similarity per se had no facilitation in the phonological processing of two-kanji compounds as in Table 8. On the other hand, semantic similarity was the main factor for naming of kanji compounds, with semantically-same kanji showing a naming latency that was 42 milliseconds faster than semantically-dissimilar kanji. Once a semantic element is added to orthographic similarity, the difference in naming latency of two-kanji compounds was amplified to 93 ms between orthographically- and semantically-same kanji and orthographically- and semantically-dissimilar kanji. Thus, with the addition of semantic similarity, orthographic similarity becomes a significant factor, even for phonological processing of kanji.

**Table 8:** Mean naming latency for orthographically and semantically similar and different compounds

| | Orthographic | Semantic | Orthographic & Semantic |
|---|---|---|---|
| Similar | 850 ms | 841 ms | 826 ms |
| Dissimilar | 873 ms | 882 ms | 919 ms |

Two types of behavioral tasks, naming and lexical decision, provide us with a clearer picture of the kanji processing mechanism. Lexical decision tasks require participants to judge whether a two-kanji compound exists as a real Japanese word. The time from the onset of visual-presentation to the judgment, indicated by pressing a YES/NO key, is measured as the reaction time. Resembling the results of the naming task, Kayamoto (2002) showed no difference between orthographically-similar kanji and orthographically-dissimilar kanji in a lexical decision task involving two-kanji compounds, as in Table 9. Again, semantic similarity was the major factor. Semantically-same kanji were processed faster than semantically-dissimilar kanji. Putting all the results of Kayamoto's studies (1995a, 1996, 2002) together, the orthographic similarity between Chinese characters and Japanese kanji has little effect on both phonological and orthographic processing of two-kanji compounds.

**Table 9:** Mean response times for orthographically and semantically similar and different compounds

| | Orthographic | Semantic |
|---|---|---|
| Similar | 645 ms | 642 ms |
| Dissimilar | 681 ms | 685 ms |

The null effects of kanji orthographic similarity reported by Kayamoto (1995a, 1996, 2002), however, may create some confusion. Her 0-to-4 scale depicting orthographic similarity is based on the measurement of a kanji unit, but the experiments were conducted on the processing of lexical kanji-compound units. Taking an example from Kayamoto (2002), an orthographically different item 階段 'stairs' (阶段 in Chinese) contains only a single orthographic difference in the kanji 階 (阶 in Chinese) which is compared against items that are orthographically identical in Japanese and Chinese such as 印刷 'print'. It is quite possible that null orthographic effects could arise from this manipulation method, in that an orthographic difference was controlled by contrasting only a single kanji in a two kanji compound. On the contrary, semantic difference/similarity was defined at the lexical level, which takes into account both characters of the compound. It is easily assumed that, since the lexical decision and naming tasks in Kayamoto (2002) involve in a combination of two kanji at the lexical level, similarity in lexical concepts naturally exerts a strong influence on lexical processing, and that lexical-level processing overrides the effects of orthographic similarity/difference at the kanji morphemic level.

Inhibitory effects of visual complexity by native speakers were found not only in Chinese characters (Leong, 1986) but also in Japanese kanji with low frequency (Tamaoka and Kiyama, 2013). As shown in Figure 2, the 1,945 kanji in the former List of Commonly-Used kanji (Jōyō kanji-hyō) have an average of 10.84 strokes with a 3.76 standard deviation (Tamaoka, Kirsner, Yanase, Miyaoka and Kawakami 2002).

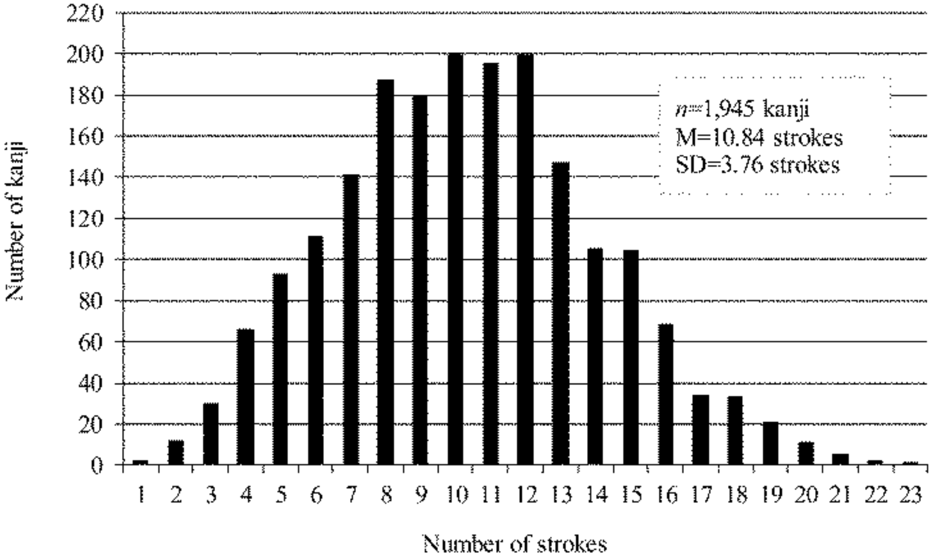



**Figure 2:** Stroke distribution of the 1,945 kanji in the former list of commonly-used kanji (Data taken from Tamaoka, Kirsner, Yanase, Miyaoka and Kawakami 2002)

Using both kanji correctness decision and kanji naming tasks, Tamaoka and Kiyama (2013) found that visual complexity inhibited the processing of low-frequency kanji among native Japanese speakers, whereas such consistency was not observed in the processing of high-frequency kanji. Kanji with medium complexity were processed faster than high-frequency simple and complex kanji. This result echoes the rather common conclusion that visually complex figures fundamentally require longer decoding times than simple ones for kanji with low frequency while high-frequency kanji display a different pattern. These studies on visual complexity were conducted under a monolingual condition with native Chinese or Japanese speakers (Leong 1986; Tamaoka and Kiyama 2013), so that effects of visual complexity and frequency on processing Japanese kanji by native Chinese speakers learning Japanese should be further re-examined in comparison with Chinese simplified characters used in mainland China and traditional complex characters used in Taiwan.

Finally, regarding the orthographically- and semantically-same words (frequently-referred as S-type words), native Chinese or Japanese speakers do not know in which language context these words are used; they have no indication whether these words are Chinese or Japanese. Following this line of reasoning, Cai and Matsumi (2009) suggested that these words are shared in the mental lexicon of both languages. This claim by Cai and Matsumi (2009) can be investigated by the following two approaches. First, these words differ in word frequencies depending on the language, therefore word frequency effects will manifest differently in the speed

of lexical processing between the two languages. In such a case, these words are separately stored in a different orthographic lexicon in each language. Second, word production size, or the number of compound words which can be produced by a single Japanese kanji or Chinese character, will differ between the two languages. Therefore, these words will behave differently depending upon the language in use. Unless these possibilities are empirically confirmed, the notion of shared word representations in a single orthographic lexicon cannot be held as a certainty.

## 3.4 Effects of kanji phonological similarity between Japanese and Chinese

Words originating from the Chinese language or created in Japan using Chinese characters often exhibit great similarity in phonology. For instance, the word 'attention' is written with two identical characters as 注 and 意 in both Japanese and Chinese. Pronunciations in both languages are very similar, spoken as /tyuu i/ in Japanese and /zhu4 yi4/ in Chinese. Like the 0-to-4 scale of orthographic similarity, Kayamoto (1995b) measured phonological similarity with a 1-to-7 point scale using subjective judgments by 11 native Chinese speakers studying at Hiroshima University whose Japanese learning experience ranged from 2 to 13 years. She used comparisons of paired kanji-pronunciations of Japanese and Chinese, such as the Japanese kanji 想 with the On-reading (a Chinese-originated sound) /soo/ in Japanese and /xiang3/ in Chinese.[3] In the actual measurement, each paired sound was presented as ソウ (/soo/) in katakana for Japanese and 'xiang' (without indication of type 3 tone) in Pinyin for Chinese. Native Chinese speakers were asked to subjectively or intuitively compare these sounds visually presented in katakana and Pinyin. In total, 1,107 pairs were presented to participants. The average rating on the 1-to-7 phonological similarity scale was 2.38 points with a standard deviation of 1.32 points, indicating that the kanji phonological similarity was rather low in its range of distribution.

Kayamoto (2000) investigated effects of phonological similarity in naming a single kanji, using a 2 × 2 design of phonologically similar and dissimilar characters between Japanese and Chinese, and Japanese On-readings and Kun-readings

3 Japanese kanji pronunciation can be divided into two types: On-readings derived from the original Chinese pronunciations, and Kun-readings originating from Japanese pronunciations (for details about kanji see Hadamitzky and Spahn 1981; Hirose 1998; Kess and Miyamoto 1999; Miller 1967; Tamaoka 1991). For example, the kanji 海 meaning 'ocean' is pronounced /kai/ in its On-reading (or Sino-Japanese) but /umi/ in its Kun-reading. On-readings are frequently used for multiple kanji compound words such as 海岸 /kaigan/ meaning 'seashore', 海賊 /kaizoku/ meaning 'pirate', and 海藻 /kaisoo/ meaning 'seaweed'. The Kun-reading frequently appears in isolated kanji, often having a concrete meaning of its own. In the case of 海, this single kanji meaning 'ocean' or 'sea' is pronounced /umi/ in the Kun-reading. On- and Kun-readings are used distinctly for different words: On-readings for *kango* (Chinese-derived words) and Kun-readings for *wago* (Japanese-based words).

(Japanese- origin sounds). She tested 12 native Japanese speakers, 12 native Chinese speakers with superior-level Japanese proficiency (or superior-level Chinese), and 12 native Chinese speakers with advanced-level Japanese (advanced-level Chinese). Phonologically similar kanji were /an/ for 案 in On-reading and /an4/ for Chinese, and /bi/ for 鼻 in On-reading and /hana/ in Kun-reading, and /bi2/ in Chinese. Phonologically dissimilar kanji were /kyoo/ for 京 in On-reading and /jing1/ in Chinese, and /tyoo/ for 鳥 in its On-reading and /tori/ in the Kun-reading, and /niao3/ in Chinese. Both advanced and superior-level native Chinese speakers named On-readings faster than Kun-readings while no difference was found among native Japanese speakers. Facilitation effects of phonological similarity were observed only among advanced-level Chinese at 79 millisecond faster in On-readings of similar and dissimilar kanji, and 50 milliseconds faster in Kun-readings of similar and dissimilar kanji, as in Table 10. Effects of phonological similarity seem to disappear, as native Chinese speakers progress in their Japanese proficiency.

**Table 10:** Mean naming latency and error rates for On- and Kun-readings

| | On reading (error rate) | Kun reading (error rate) |
|---|---|---|
| Similar | 787 ms (10.3%) | 860 ms (17.3%) |
| Dissimilar | 866 ms (12.2%) | 910 ms (10.3%) |

It should be noted, however, that phonological similarity measured by Kayamoto (1995b) is defined based on On-readings, not Kun-readings. Phonological similarity does not refer to a similarity index for Kun-readings. Furthermore, kanji with Kun-readings are always accompanied with On-readings in Kayamoto (2000). Since multiple readings, including both On- and Kun-readings are activated when native Japanese speakers encounter kanji (Verdonschot, La Heij, Tamaoka, Kiyama, You and Schiller 2013), native Chinese speakers must puzzle over which On-reading or Kun-reading they should chose to pronounce. A delay in Kun-reading could be a result of this selection process amongst multiple phonological activations of a single kanji such as /zi/ in the On-reading and /mimi/ in the Kun-reading for 耳, or /seki/ in the On-reading and /aka/ in the Kun-reading for 赤. Therefore, Kayamoto (2000)'s conclusion must be limited to only kanji with On-readings, but not to kanji with Kun-readings as such: Advanced-level Chinese had facilitation effects of phonological similarity on phonological processing of On-readings, and these disappear once they reach a higher level of Japanese proficiency.

Kayamoto (2002) also investigated phonological similarity effects on naming compound words constructed with two On-reading kanji (e.g., 感謝 /kansya/, 銀行 /ginkoo/ and 無心 /musin/). Results indicated facilitation effects among native Chinese speakers learning Japanese such that phonologically-similar two-kanji compounds (M = 829 ms) were named faster than phonologically-dissimilar ones

(M = 893 ms), while no difference was found in the lexical decision task. Unlike Kayamoto (2000), Kayamoto (2002) used a naming task involving two-kanji compound words which usually have only a single reading. Thus, it is safe to conclude that phonological similarity facilitates naming speed for two-kanji compound words with a combination of On-readings.

## 3.5 Semantic similarities and differences in kanji compound words between Japanese and Chinese

The Agency for Cultural Affairs in Japan (1978) provided a lexical typology of Japanese kanji-compound words corresponding to Chinese words. The agency classified kanji compounds into four types. (1) Same-type (S-type) refers to the same meaning between Japanese and Chinese. Two-thirds of all kanji-compound words are classified into this S-type, e.g., *ondo* 温度 'temperature' and *mirai* 未来 'future'. Native Chinese speakers indeed have a great advantage learning Japanese vocabulary of S-type. (2) Overlapping-type (O-type) is defined as meanings partly overlapped between the two languages. Words in O-type have intricate interactions between the two languages, e.g., *binboo* 貧乏 'poverty', and *hakusi* 白紙 'white paper' or 'annul'. (3) Different-type (D-type) implies kanji-compound words semantically different from their Chinese meanings, e.g., *tegami* 手紙 'letter' in Japanese but 'toilet paper' in Chinese, and *monku* 文句 'complain' in Japanese but 'sentence and phrase' in Chinese. (4) Nothing-type (N-type) implies no corresponding words ('nothing') exist in Chinese, e.g., *taikutu* 退屈 'boredom' and *okubyoo* 臆病 'timidity'. Previous studies (e.g., H. Chiu 2002, 2003; Y. Chiu 2006, 2007; Hayakawa 2010; Hayakawa and Tamaoka 2012) conducted experiments on lexical processing by comparing S-type and N-type words (see details, Sections 3.5). The meanings of some N-type Japanese words are easier to guess from the knowledge of Chinese characters, but some are not. It is possible to classify a fifth type of words or kanji combinations which only exist in Chinese, such as 公司 /gong1 si1/ 'company', but this category is unnecessary for the purpose of comparing Japanese and Chinese.

Komori and Tamaoka (2010) classified O-type compounds into three sub-categories as shown in Figure 3 (i) those with meanings particular to Chinese, (ii) those with meanings particular to Japanese, and (iii) those with meanings particular to both Japanese and Chinese.

The first sub-category of (i) O-type Sub-1 is defined as kanji-compound words that partly share the same meaning(s) in both Japanese and Chinese, but for which Chinese contains its own extended meanings. For example, 貧乏 /bin boo/ 'poor' in Japanese can be used as in 'poor life' expressed as *binboo-na seikatu* 貧乏な生活 in Japanese, and *pinfa shenghuo* 贫乏生活 in Chinese. The meaning of this word is extended to use with 'experience' as 'poor experience' *pinfa jingyan* 贫乏经验 in Chinese, but not in Japanese. Likewise, 贫乏 in Chinese can be used with 'thinking',

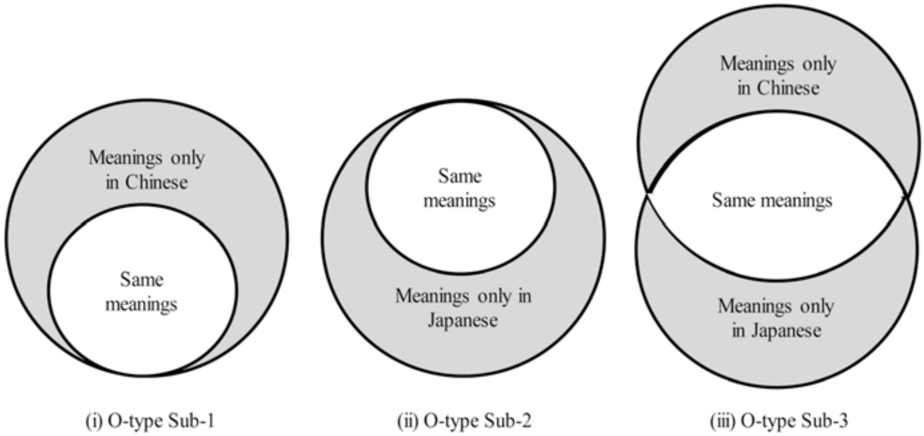



**Figure 3:** Sub-categories of overlapping-type (O-type) kanji compound words (The figure is from Komori and Tamaoka (2010: 166) with partial modification.)

'thought', and 'idea' as in *sixiang pinfa* 思想贫乏 'poor in thought'. Because of these extended meanings in Chinese, native Chinese speakers are likely to overextend the usage of this word to produce incorrect Japanese expressions such as *keiken-ga binboo-da* 経験が貧乏だ '(my/your) experience is poor', and *kangaekata-ga binboo-da* 考え方が貧乏だ '(my/your) way of thinking is poor'.

The second sub-category of (ii) O-type Sub-2 is defined as those words partly sharing the same-meaning(s) in both Japanese and Chinese, but featuring extended meanings in Japanese. The word 貴重 /ki tyoo/ meaning 'valuable' or 'precious', for example, can be used as 貴重品 *kityoo-hin* 'a valuable article' in Japanese and as *guizhong-pin* 贵重品 in Chinese. The word is used with 'time' as in *kityoo-na zikan* 貴重な時間 'valuable time' and with 'experience' as in *kityoo-na keiken* 貴重な経験 'valuable experience' in Japanese, but not in Chinese. Instead, 宝贵 /bao3 gui4/ is used in Chinese as in *baogui shijian* 宝贵时间 'valuable time', and as *baogui jingyan* 宝贵经验 'valuable experience'. Because of these differences in usages, it is difficult to acquire expressions containing O-type Sub-2 words like *kityoo-na zikan-o saite* 貴重な時間を割いて 'to spare valuable time'. Yet, if native Chinese speakers avoid using these Japanese expressions which are not found in Chinese, they will not make mistakes.

The third sub-category of (iii) O-type Sub-3 is defined as those words which partly share the same meaning(s) in both Japanese and Chinese, but which contain extended meanings both in Japanese and Chinese. For example, 是非 pronounced /ze hi/ in Japanese and /shi4 fei1/ in Chinese has multiple meanings. In both Japanese and Chinese, this word can be used with the meaning of 'right or wrong' as in the expression *zehi-no kubetsu-ga aimai-da* 是非の区別があいまいだ 'A distinction of right and wrong is unclear' in Japanese, and *bu fen shifei* 不分是非

in Chinese. This word can also be used differently in Japanese and Chinese. In Japanese, this word is used to mean ‘please’ in *zehi go-sanka kudasai* 是非ご参加ください ‘Please participate in it’, but there is no such usage in Chinese. Contrarily, this word is used to mean ‘a quarrel’ in Chinese as in *re shifei* 惹是非 ‘Picking quarrels’, but has no such meaning in Japanese.

Komori and Tamaoka (2010) investigated how native Chinese speakers learning Japanese process words of O-type Sub 1 and O-type Sub 2. Using their original Cloze Test, they selected 22 Chinese with higher-level Japanese proficiency (M = 71.45, SD = 3.88) and 22 Chinese with lower-level Japanese proficiency (M = 44.68, SD = 4.31) from 64 participants studying in Japan. The Cloze Test required the participants to fill in the missing words removed from a text. They obtained a very high Cronbach’s alpha reliability of 0.946 ($n$ = 64, M = 58.09, SD = 11.96). A priming experiment was then conducted in which a priming word was presented for 280 milliseconds, and a target word presented following a 120 millisecond interval. The interval between the prime onset and the target onset times (i.e., stimulus-onset asynchrony, SOA) was 400 milliseconds. In Experiment 1 of the processing of O-type Sub-1 words, they conducted the Chinese lexical decision task under the priming condition. The results showed that primed Chinese words of both shared meanings (e.g., 方位 ‘direction’) and the meanings particular to Chinese (e.g., 物品 ‘commodity’) significantly facilitated the lexical decision times of the target Chinese words (e.g., 东西) to the same degree regardless of the level of Japanese proficiency. For the Japanese lexical decision task in Experiment 2, which required processing of O-type Sub-2 words among Chinese speakers with high Japanese proficiency, primed Japanese words with shared meanings (e.g., 細心 ‘scrupulous’) facilitated the lexical decision times of the target Japanese words (e.g., 注意), but words with meanings unique to Japanese did not (e.g., 警告 ‘warning’). By contrast, among Chinese with lower Japanese proficiency, neither primed words of shared meaning nor those unique to Japanese facilitated processing of the target Japanese words.

The results of priming effects in Experiment 1 (L1 Chinese condition) of Komori and Tamaoka (2010) suggest that orthography and concepts were very strongly linked in the Chinese mental lexicon. However, null priming effects found among native Chinese speakers with lower Japanese proficiency in Experiment 2 (L2 Japanese condition) indicate smaller and weaker connections from orthography to concepts in the Japanese mental lexicon. This contrasting finding further suggests that the size and strength of lexical connections in L2 Japanese between orthography and concepts are less robust than those of the first language (Chinese). Yet, the fact that priming effects of the shared meanings were apparent among those with higher Japanese proficiency indicates that the higher the proficiency level they reach in their second language, the stronger the connections between lexical and conceptual representations become in their second language. However, since null priming effects were found for words with the Japanese-particular meanings, it seems that the Japanese-unique meanings are difficult for native Chinese speakers to acquire.

As such, the difficulty with the Japanese-particular meanings and usages among O-type Sub 2 and Sub 3 words are revealed in the priming study.

## 3.6 Differences in On- and Kun-readings in kanji phonological processing

Using the index of kanji On-reading ratios calculated by Kaiho and Nomura (1983), Tamaoka and Taft (2010) reported that kanji with a 50 percent On-reading ratio randomly embedded with kanji in an On-reading dominant environment were mostly pronounced in On-readings; likewise, the same target kanji embedded with kanji in a Kun-reading dominant environment were mostly pronounced in Kun-readings. Native Japanese speakers easily shifted between On- and Kun-readings, depending on the phonological context. That is, separate On- and Kun-reading sub-lexica exist within the phonological lexicon.

If native Chinese speakers have a well-established sub-lexicon of On-readings associated with characters and their compound words in L1 Chinese, they can produce On-readings faster than Kun-readings. In fact, H. Chiu (2003) showed that kanji compounds with On-readings were named faster than those with Kun-readings among native Chinese speakers who had attained the first and second level of the Japanese Proficiency Test. Thus, native Chinese speakers are likely to associate phonology in Chinese to On-readings more easily than to Kun-readings. A question arises whether phonological suppression by inter-lexical interference for cognates (Hayakawa 2010; Hayakawa and Tamaoka 2012) conflicts with the advantage of On-readings over Kun-readings. Kun-readings are fundamentally used for non-cognates, and the number of kanji compounds with Kun-readings (*wago*) is much smaller than those with On-readings (*kango*). Because On-readings are associated with both cognates and non-cognates, the advantage of On-readings and the phonological suppression for cognates should be treated as a separate issue.

## 3.7 Lexical processing differences for cognates and non-cognates

The term *cognate* is often used in bilingual studies on languages spoken in Europe. In linguistics, this term refers to words of a common etymological origin. A typical example of a cognate in Indo-European languages is the word *night* in English. Spelling or orthography of this word differs depending on the language, as in French *nuit* and German *Nacht*, and the Dutch *nacht*, with the same spelling as German. In psycholinguistics, cognates are denoted as words similar in orthography, phonology, and semantics. Thus, cognates described in linguistics do not totally overlap with those in psycholinguistics. When explaining studies on kanji processing in this section, I will follow the psycholinguistic definition, ignoring the etymological connotation of the term.

Bilingual studies on European languages have clearly indicated that cognates (similar in spelling, sound, and meaning) are processed faster than non-cognates (e.g., Costa, Caramazza and Sebastián-Gallés 2000; de Groot, Delmaar and Lupker 2000; Dijkstra and van Heuven 2002; Green 1998; van Heuven, Dijkstra and Grainger 1998; van Heuven, Schriefers, Dijkstra and Hagoort 2008). Cognates for kanji-compound words between Chinese and Japanese are defined as orthographically-similar and semantically-same words. For example, the Japanese two-kanji compound word 法則 is represented as the two orthographically-similar kanji 法则 in Chinese, having the same meaning 'law'. This word is pronounced quite differently /hoo soku/ in Japanese, and /fa3 ze2/ in Chinese though; consequently, the term cognate does not refer to phonological similarity. Conversely, the term *non-cognates* is defined as orthographically- and semantically-different words. An example is 財布 'wallet' /sai hu/ in Japanese. This combination of two kanji does not exist in Chinese, with 'wallet' being 钱包 /qian2 bao1/ in Chinese. This Chinese word can be written using two orthographically-similar kanji 銭包 in Japanese, which, of course, does not exist in Japanese. Since a majority of kanji are basically shared in both languages, the real difference between cognates and non-cognates among kanji-compound words is the way in which kanji are combined.

A unique difference was found between cognates and non-cognates in processing Japanese kanji-compound words by native Chinese speakers. H. Chiu (2003) conducted a naming experiment on three different types of words: cognates, non-cognates with On-readings, and non-cognates with Kun-readings. The experiment was conducted with four different groups; native Chinese speakers (studying Japanese at a university in Taiwan) with the second ($n$ = 17) and first ($n$ = 19) levels of the Japanese Proficiency Test, those with highly advanced Japanese ($n$ = 15) studying in Japan, and also native Japanese speakers ($n$ = 20). She controlled participants' age of acquisition (AoA) of kanji-compound words. AoA is defined as the age at which a word is learned in acquiring spoken language. Morrison and Ellis (1995) found a strong AoA effect when word frequency was controlled, but no word frequency effect when AoA was controlled. The stimulus manipulation of AoA by H. Chiu (2003), however, differs from Morrison and Ellis (1995). She divided the stimuli into two groups – beginner level for Japanese words with an early AoA and intermediate level for words with a late AoA – based on difficulty-levels of words provided by the Japan Foundation and Japan Educational Exchange and Services (2002). In order to avoid confusion with AoA studies in English, in which some studies have only reported minor effects (e.g., Zevin and Seidenberg 2002) in contrast with Morrison and Ellis (1995), I describe early AoA as easy words and late AoA as difficult words in the following explanation of Chiu's results.

H. Chiu (2003) found unexpected results between cognates and non-cognates in her naming task. Native Chinese speakers of the second level (intermediate Japanese) showed a trend in ascending order of naming latencies on easy words; non-cognates with On-readings, cognates, and non-cognates with Kun-readings as in Table 11.

This trend was much clearer among difficult words, with an ascending order of non-cognates with On-readings, cognates, and non-cognates with Kun-readings.

**Table 11:** Mean naming latency for cognates and non-cognates in *On*- and *Kun*-readings

| | | On reading | | Kun reading |
|---|---|---|---|---|
| | Words | Non-cognates | Cognates | Non-cognates |
| 2nd Level (Intermediate) | Easy | 846 ms | 893 ms | 1,151 ms |
| | Difficult | 948 ms | 1,135 ms | 1,232 ms |
| 1st Level (Advanced) | Easy | 852 ms | 806 ms | 991 ms |
| | Difficult | 875 ms | 948 ms | 987 ms |

In comparison, Chinese learners of the first level (advanced Japanese) displayed no difference between cognates and non-cognates with On-readings on easy words. However, cognates were named faster than non-cognates with Kun-readings. Among difficult words, the previous trend was observed again in the ascending order of non-cognates with On-readings, cognates, and non-cognates with Kun-readings. This trend was observed neither among Chinese-speaking learners with highly-advanced Japanese, nor among native Japanese speakers. Error rates also indicated a very similar overall tendency.

Unlike the facilitation effects of cognates in European languages (e.g., Costa, Caramazza and Sebastián-Gallés 2000; de Groot, Delmaar and Lupker 2000; Dijkstra and van Heuven 2002; Green 1998; van Heuven, Dijkstra and Grainger 1998; van Heuven, Schriefers, Dijkstra and Hagoort 2008), inhibitory effects were found in H. Chiu (2003) demonstrating that cognates were named much more slowly than non-cognates among native Chinese speakers who were less proficient (both easy and difficult words) and who were advanced speakers in Japanese (only difficult words). Based on this result, H. Chiu (2002, 2003) proposed that the processing routes of kanji-compound words varies depending on the lexical relationship between Japanese and Chinese, and she constructed a phonological processing model which contrasts cognates and non-cognates. As depicted in (i) of Figure 4, cognates are first processed through the phonological route in Chinese, and then further to their concepts. In contrast, as shown in (ii) of Figure 4 non-cognates do not exist as Chinese words, so that newly-acquired non-cognates are easily processed through the Japanese sound route. Due to the difference of these two phonological processing routes, the naming of cognates in Japanese is slowed down, whereas non-cognates are pronounced more quickly in Japanese than cognates.

The model in Figure 4 (H. Chiu 2002, 2003) was supported by several studies (Y. Chiu 2006; Hayakawa 2010; Hayakawa and Tamaoka 2012; Komori 2005) and partly by a paper by Y. Chiu (2007). Komori (2005), as previously described, showed that Chinese speakers have an advantage in reading comprehension, but not in

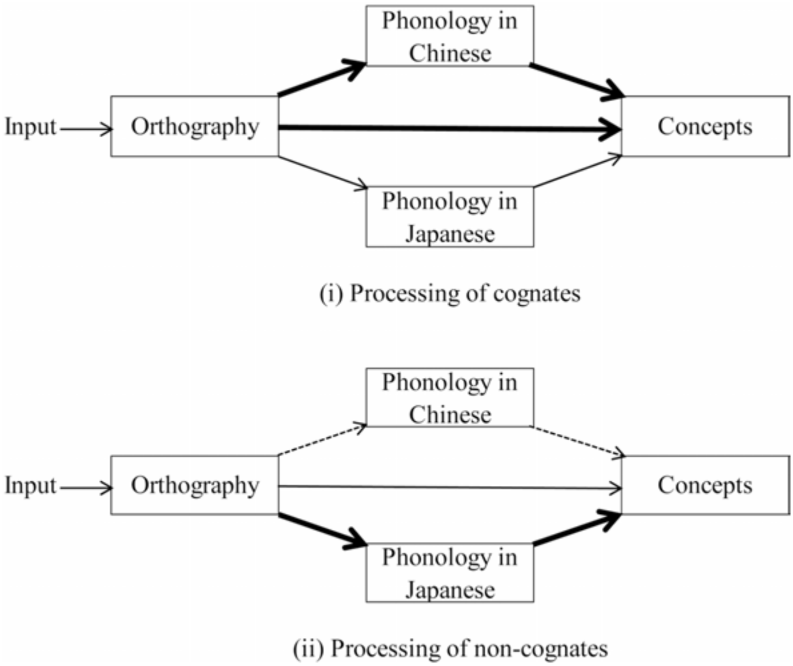



**Figure 4:** Difference in phonological processing of cognates and non-cognates by native Chinese speakers learning Japanese (The figure is taken from H. Chiu (2002) and translated into English.)

listening comprehension due to their kanji orthographic knowledge. This study further described that cognates (56.35% accuracy) were understood nearly as well as non-cognates (51.18% accuracy) in listening comprehension, while cognates (80.57% accuracy) yielded a greater advantage than non-cognates (64.07% accuracy) in reading comprehension. Y. Chiu (2006) conducted an experiment with 12 native Chinese speakers who had passed the first level of the Japanese Proficiency Test, employing a lexical decision task for words embedded in a sentence. In her study, a sentence containing parentheses as in ( ) で勉強したので，とても眠い 'Because I studied ( ), I am very sleepy' was visually presented, followed by the auditory presentation of a compound word. In this sentence, a possible correct response is the word /tetuya/ 'all night'. Participants were required to decide whether this word appropriately fits into the parentheses in the sentence. Kanji compound words which were cognates required more time for participants to make a lexical decision than non-cognates. Since stimulus words were auditorily presented, as shown in Figure 4, non-cognates must be strongly tied to both Japanese phonology and concepts, while cognates must have only loose ties with Japanese phonology.

Hayakawa (2010) and Hayakawa and Tamaoka (2012) provided support for the model in Figure 4. Hayakawa (2010) tested 48 Chinese speaking learners of Japanese

(26 at the first level and 22 at the second level of the Japanese Proficiency Test). In order to investigate the effects of traditional Chinese characters, which are used primarily in Taiwan, she selected kanji compounds based on orthographic figures of traditional characters. For the lexical decision task using auditory-presented words, Hayakawa (2010) chose three different types of 16 kanji compound words each (48 target words in total): (1) S-type (e.g., 記憶 in Taiwan and Japan) – orthographically-/semantically-same compounds that are considered to be cognates in H. Chiu (2003), (2) D-type (e.g., 作業 in Japan 工作 in Taiwan) – orthographically-similar but semantically-different, and (3) N-type (e.g., 退屈 only used in Japan) – two-kanji combinations that do not exist in Chinese and are considered non-cognates in H. Chiu (2003).

**Table 12:** Mean response times for auditory-presented cognates and non-cognates

| | S-Type | D-Type | N-Type |
|---|---|---|---|
| 2nd Level (Intermediate) | 1,400 ms | 1,283 ms | 1,192 ms |
| 1st Level (Advanced) | 1,201 ms | 1,143 ms | 1,086 ms |

Although H. Chiu (2003) did not obtain a clear trend among Chinese learners of advanced Japanese on difficult words (or late AoA), Hayakawa found an ascending order of both N-type, D-type, and S-type among the Intermediate level learners ($n$ = 22), and N-type, D-type, and S-type among the advanced learners ($n$ = 26), as in Table 12. With this result, phonological inhibitory effects for cognates were extended to a wider population of native Chinese speakers including those at the advanced learners or those who passed the first level of the Japanese Proficiency Test.

Furthermore, Hayakawa and Tamaoka (2012) examined phonological processing of S-type (cognates) and N-type (non-cognates) in lexical decisions of auditory-presented words, using 38 native Chinese speakers from mainland China and 38 Korean speakers (control group) learning Japanese. Once again, lexical decisions were slower for S-type than N-type among Chinese, whereas no difference was found between S-type and N-type among Koreans, as in Table 13. Since Koreans use little of the kanji script in their language, and since S-type and N-type were classified based on similarity in Chinese character words, null effects among the Korean control group strengthened the results found with the Chinese participants.

**Table 13:** Mean response times for auditory-presented cognates and non-cognates

| | S-Type | N-Type |
|---|---|---|
| Chinese ($n$ = 38) | 1,188 ms | 1,111 ms |
| Korean ($n$ = 38) | 1,168 ms | 1,157 ms |

Hayakawa (2010) and Hayakawa and Tamaoka (2012) explained the processing mechanism in detail as follows. Cognates of kanji compounds already have phonological representations in the Chinese mental lexicon. For instance, 未来 is pronounced /wei4 lai2/ in Chinese. To acquire this word in Japanese, a native Chinese speaker has to memorize its Japanese sound /mi rai/ in addition to their prior knowledge of the Chinese sound. In doing so, the orthography of the cognate 未来 becomes simultaneously connected to two different phonological representations, /wei4 lai2/ in Chinese and /mi rai/ in Japanese. The connection from orthography (未来) to phonology (/wei4 lai2/) in the first language is very strong, but the newly-learned sound of the word (/mi rai/) has a relatively weak connection. As a result, when the cognate is presented auditorily, the newly-learned Japanese phonology /mi rai/ delays the activation of the Chinese pronunciation in reaching its necessary threshold. On the other hand, since there is no lexical phonology in Chinese for non-cognates, the newly-learned sound of a non-cognate is easily activated without competition from existing Chinese phonological representations.

## 4 Processing syntactically different features

Are native Chinese speakers learning Japanese unable to break free from the spell of Chinese syntactic features when processing the Japanese language? Due to the great syntactic difference between Japanese and Chinese, or so-called *longer linguistic distance* in syntax, it is frequently presumed that native Chinese speakers have greater difficulties in processing Japanese sentences compared to native Korean speakers whose language, in terms of syntax, is considered to exhibit *shorter linguistic distance* (Horiba and Matsumoto 2008; Koda 1993, 2005). However, according to Fan and Wu (2006), among second-year native Chinese speakers majoring in the Japanese language at Xi'an International Studies University, 79.79% in 2002 and 82.26% in 2003 passed the fourth level of the Japanese language specialization test (i.e., Nihongo Senmon Shiken 4; NSS4) conducted by the Ministry of Education in the People's Republic of China. The fourth level is said to be equivalent to the second level of the newer Japanese language proficiency test (i.e., N2) administered by the Japan Foundation. Furthermore, although there is no specific data available, it is commonly known among instructors of the Japanese language in China that approximately half of the native Chinese speakers majoring in Japanese at the eight major universities of foreign languages in China (i.e., two in Beijing, and one in Dalian, Guangzhou, Shanghai, Sichuan, Tianjin, and Xi'an) can pass the highest level of the Japanese language proficiency test (i.e., N1) at the end of three years of Japanese study, even when starting with no Japanese knowledge. Given this remarkable improvement in such a short period of learning Japanese, it may not be difficult for Chinese students to overcome syntactic differences between Chinese and Japanese to the degree that researchers have previously assumed.

## 4.1 Morphosyntactic inflections and differences in Chinese and Japanese

Many learners of Japanese with no kanji background devote numerous hours to memorizing kanji orthography. By contrast, thanks to the high degree of orthographic similarity between Japanese kanji and Chinese characters, native Chinese speakers can allocate the majority of their classroom and study hours to learning Japanese grammar or syntactic features from the early stages of study. If so, despite the Chinese language being syntactically dissimilar to Japanese (i.e., a longer linguistic distance in syntax), high levels of achievement would be expected in acquiring Japanese grammar over a short period among Chinese students. This implies a strong version of the predicted learning potential of native Chinese speakers in that they will likely encounter little syntactic difficulty in learning Japanese.

The Chinese language has no morphosyntactic inflections. This poverty of syntactic features is expected to present difficulties in the acquisition of Japanese verb inflections (if one assumes L1 transfer). Chu, Tamaoka and Yamato (2012) investigated how 102 native Chinese speakers learning Japanese acquire verb inflections during only a four month period at a university in China. Participants were tested on *te*-form verb inflections which were reported as being very difficult for Japanese learners (e.g., Nagatomo 1997; Sakamoto 1993). Cronbach's reliability for 54 target verbs by Chu et al. was very high at $\alpha$ = 0.86. Their verbs were taken from four sources, (i) 15 verbs from the students' textbook (e.g., *oyogu* 'swim' and *au* 'meet'), (ii) 15 verbs not in the textbook (e.g., *mayou* 'lost' and *susumu* 'progress'), (iii) 15 nonsense verbs created by the authors (e.g., *kaziku* and *miaru*), and (iv) 9 recently-coined verbs (e.g., *tikuru* 'secretly tell someone' and *kokuru* 'confess one's feelings'). All students were asked to write the *te*-form inflections of all 54 verbs. For example, when *yomu* 'read' is presented, participants must write the correct *te*-form *yonde* for 1 point.

The results of Chu et al. (2012) are indicated below in descending order of accuracy; 90.33% for verbs from the textbook < 88.00% for nonsense verbs = 87.20% for verbs not in the textbook = 86.44% for newly-created verbs. Verbs taken from the textbook and used in their classroom were better than those from other categories. What is more surprising is that Chinese students exhibited over 86% accuracy on all four categories. They further reported difficulty levels of *te*-inflections depending on forms, indicated in descending order of accuracies; –*tte* form (98.13%, e.g., *atte* 'meeting') > –*site* form (94.04%, e.g., *zyunbisite* 'preparing') = –*te* form (90.76%, e.g., *mite* 'seeing') > –*ite/ide* form (85.59%, e.g., *oyoide* 'swiming') > -*nde* form (74.03%, e.g., *susunde* 'progressing'). Besides the –*nde* form, all other forms displayed high performance at over 85% accuracy.

It is amazing that native Chinese speakers could apply the *te*-form inflection rules to nonsense verbs after a mere four months of Japanese study; a difference in accuracy of only 2.33% between verbs in the textbook and nonsense verbs (90.33%–

88.00%). Proper application of inflectional morphology for nonsense verbs is considered an indication of well-formed, rule-based knowledge. According to these results, native Chinese speakers learning Japanese, even for a period of only four months, adequately apply their acquired knowledge of *te*-inflection rules to various verbs, despite the absence of inflectional morphology in their first language. In this sense, language acquisition researchers generally seem to be overestimating the negative effects of linguistic differences in syntax between Chinese and Japanese. The absence of syntactic features may not be a crucial obstacle for acquiring Japanese, although it provides no facilitation. It should, however, be noted that Chu et al. (2012) simply asked native Chinese speakers to inflect a verb stem. They did not test the actual use of verbs in a sentence. Therefore, acquisition of verbal inflections should be further investigated by means of on-line processing of a sentence predicate.

Difficulties in processing two-kanji compounds by native Chinese speakers could be found when noun compounds are used as verbs (i.e., verbal nouns). Native Chinese speakers are likely to apply their knowledge of Chinese to Japanese, even though some verbal nouns differ in their usage, such as transitive/intransitive and active/passive. For example, as shown in examples (6a) and (6b), a majority of verbal nouns are used for active and passive in both Japanese and Chinese (e.g., *kakunin* 'check'). Yet, some verbal nouns are used in active form in both Japanese and Chinese, but with the passive used only in Japanese (e.g., *zyunbi* 'prepare') as shown in (7a) and (7b).

(6) a. Active form used in both Chinese and Japanese

| | | |
|---|---|---|
| *Suuti o* | *aratani* | *kakuninsita.* |
| numerical value ACC | newly | check PST |

'(He) newly checked the numerical values.'

b. Passive form used in both Chinese and Japanese

| | | |
|---|---|---|
| *Suuti ga* | *aratani* | *kakuninsareta.* |
| numerical value NOM | newly | check PASS PST |

'The numerical values were newly checked.'

(7) a. Active form used in both Chinese and Japanese

| | | |
|---|---|---|
| *Siryoo o* | *keikakutekini* | *zyunbisita.* |
| reference ACC | deliberately | prepare PST |

'(He) deliberately prepared the reference.'

b. Passive form used only in Japanese

| | | |
|---|---|---|
| *Siryoo ga* | *keikakutekini* | *zyunbisareta.* |
| reference ACC | deliberately | prepare PASS PST |

'The reference was deliberately prepared.'

A majority of two-kanji compound nouns shared by Chinese and Japanese are fundamentally used in the same way as shown in (6a) and (6b). As a result, native Chinese speakers are predicted to show no qualitative differences among sentence types in (6a), (7a), and (6b). However, if the morphosyntactic knowledge of Chinese words is merely applied to Japanese, lower accuracy and possibly slower speed are expected to occur in the processing of passive sentences like in (7b), which is not used in Chinese. This type of subtle difference observed in verbal nouns is expected to lead to occasional, but unavoidable mistakes. Morphosyntactic knowledge of Chinese will therefore likely cause considerable influence on the processing of second language Japanese two-kanji compounds. The question of on-line predicate processing by Chinese speakers still remains to be answered in future studies.

## 4.2 Word order and processing Japanese sentences

Japanese base word order is SOV while Chinese one is SVO. Due to the different word order in their L1 and the target language, Chinese speaking learners of Japanese may face difficulty in the processing of even simple Japanese sentences because of the different base word order and flexible word order, i.e., scrambling (for processing of Japanese scrambled sentences, see Koizumi's and Chang's chapters in this volume.) They are required not only to process SOV-ordered Japanese sentences, but also to comprehend sentences with a phrasal movement operation of OSV scrambled order. According to word order typology by Dryer (2012), SVO and SOV are the two major types, with 41.08% being SVO (488 languages) and 47.56% being SOV (565 languages) out of 1,188 languages (a total of 1,377 minus 189 languages lacking a dominant word order). Both the Chinese and Japanese languages are included in the two major language typologies.

An early cross-linguistic study by Koda (1993) measured sentence correctness among Chinese, English, and Korean speaking learners of Japanese at an American university. Note that her study indexes the end result of sentence processing because she did not measure the reaction time of each sentence. The study showed the null effect on scrambling by Koreans. This result must have been caused by a measurement limitation in which the Korean speakers had reached the performance ceiling in terms of comprehension of the 12 total sentence stimuli in both canonical (M = 11.5) and scrambled (M = 12.0) order under the condition where case particles were present. Without case particles, however, they seemed to lose cues for processing, resulting in lower scores for both canonical (M = 8.5) and scrambled (M = 8.6) sentences, though there was still no scrambling effect. Here it should be noted that sentences without case particles are considered to be incorrect in Japanese, so that it

is problematic to estimate the mechanism for those sentences processed by any of the three language groups.[4]

In contrast with Koreans, the scrambling effect was apparent for both the American (native English speakers) and Chinese groups (see Koda 1993, Table 1). Koda drew the rather unclear conclusion that Japanese sentence processing of canonical and scrambled orders by L2 learners involves both L1 and L2 effects. It is tempting to interpret these results in such a way that American and Chinese learners were able to establish a filler-gap dependency (the relationship between the moved landing site and the original position where it was moved from) for processing scrambled sentences in a similar way to native Japanese speakers, yielding lower accuracy in the scrambled condition. However, the interpretation of gap-filling parsing is a great logical jump to apply to the results since Koreans, whose first language has case particles similar to Japanese, did not show the scrambling effect.

Selecting participants from students at an American university invites two major potential weaknesses. First, it is difficult to know how efficiently these students can handle their first languages of Chinese and Korean, as their length of residence in the US was unknown. Second, all participants may have a great deal of variation of proficiency in English. Some of the Chinese speakers may no longer have Chinese as their dominant language; instead, English may have become the more highly activated of their two languages. So-called ‘heritage’ learners who grew up in the US with Chinese/Korean parents are likely to be more English dominant. On the other hand, for students who arrived in the US after attending high school in their home country, Chinese/Korean usually remains their dominant language. With potentially low Chinese ability, can we still say these participants are good representatives of native Chinese speakers? In contrast, as far as they can proficiently speak the Korean language, Koreans may have the advantage of speaking an SOV-ordered language because Japanese also has the same SOV-order. In fact, the ceiling score in Koda’s study may be the result of syntactic similarity between Japanese and Korean. Nevertheless, both the Chinese and Korean participants must have already obtained an excellent level of English ability as their second or possibly first language, and thus, Japanese must necessarily be the third language for them. The effect from their English knowledge remains unknown.

In addition to accuracies on sentence correctness decisions, Koda conducted a reading comprehension test. A regression analysis showed that case particle knowledge ($R^2 = 0.4795$) was a highly significant predictor of reading comprehension ($p < .0001$).[5] This result clearly established a causal relation between the knowledge

---

**4** Koda (1993) was testing the strengths of different cues (e.g., animacy, case particles, word order) in the competition model (Bates and MacWhinney 1987). Thus, she used the unnatural sentences. See Shirai’s chapter in this volume on the competition model.

**5** In her article, the $R^2$ value of 47.95 is printed in Table IV, which, I assume, must have a mistake in the decimal point.

of case particles and reading comprehension. However, because she did not report group differences on the reading comprehension scores, the question to be raised is whether the Korean participants were higher achievers than the American and Chinese participants at the time when they were tested. Future cross-linguistic studies should be conducted by controlling the Japanese ability of different first-language groups, ideally focusing on Japanese being learned in the second language environment, not a foreign language environment.

Experimental approaches measuring reaction times are rather scarce in the study of Japanese sentence processing by Chinese speaking learners of Japanese. One of the few examples is Tamaoka (2005, Experiment 1) which investigated how Chinese learners who studied Japanese for two to four years at a university in Dalian, China, processed and made correctness decisions on active transitive-verb sentences with canonical and scrambled orders. Because sentence processing requires a heavy cognitive load, Tamaoka selected 24 participants out of 87 native Chinese speakers with scores higher than 22 points or 91.7% accuracy based on the results of a grammar test with 25 multiple choice questions (i.e., a maximum score of 25).

The results showed that simple active sentences in canonical order were more quickly and accurately processed than the same sentences in scrambled order as in Table 14.

**Table 14:** Mean response times and accuracy for canonical and scrambled sentences

| | Response times | Accuracy |
|---|---|---|
| Canonical | 3,566 ms | 87.5% |
| Scrambled | 3,933 ms | 78.0% |

A scrambling effect of 367 milliseconds in reaction time and 9.5% accuracy suggests the possibility that it is highly probable that the Chinese participants generated the base structure $[_{S}$ NP$_{\text{-NOM}}$ $[_{VP}$ NP$_{\text{-ACC}}$ V]] for active transitive-verb sentences and established a filler-gap dependency for scrambled-ordered sentences, as with native Japanese speakers (Aoshima, Phillips and Weinberg 2002; Koizumi and Tamaoka 2004, 2006, 2010; Mazuka, Itoh and Kondo 2002; Miyamoto and Takahashi 2002; Sakamoto 2002; Tamaoka et al. 2005).

Native Chinese speakers learning Japanese understood simple Japanese active sentences with a transitive verb in the SOV-canonical order like ‘My elder sister ate an apple’ more quickly and accurately than OSV-scrambled orders. This provides evidence that they may manipulate syntactic operation for the scrambled order. If so, at least, this finding does not support the *shallow structure hypothesis* proposed by Clahsen and Felser (2006), which claims that second language learners can process semantic roles such as lexical items, but not syntactic information, even at the

advanced level. Rather, the result supports the claim by White (2003) that syntactic features related to functional categories could be acquired in an early stage of second language acquisition, although some features such as the definiteness of determiners *a* and *the* in English are very difficult for Japanese and Chinese speaking L2 learners to acquire because their L1 lacks such a feature (Trenkic 2002).

The Chinese language has no overt *wh*-movement. *Wh*-words stay *in situ* in Chinese (He 2000; Huang 1981; Lin 1998). For example, English sentence (8a) is expressed as (8b) in Chinese. Likewise, (9a) is expressed as (9b).

(8) a. *What do you eat?*

b. 你吃什么?
*ni3 chi1 shen2-me0.*
you NOM eat PRS what ACC

(9) a. *Whom do you like?*

b. 你喜欢谁?
*ni3 xi3-huan0 shui2.*
you NOM like PRS whom ACC

Syntactic operations of English *wh*-questions include an additional fronting operation of a *wh*-word, compared to yes/no-questions that require the insertion of *do* when verbs are regular (not BE). Yet, adult native Chinese speakers, who have been studying at universities where English is the instructional language, seem to be able to handle *wh*-questions in English fairly well. Thus, it is anticipated that they can also process scrambled sentences in Japanese using a filler-gap parsing operation.

Experiment 2 in Tamaoka (2005) further examined potential sentences whose case particles conflicted with the grammatical information of subject and object taken from the stimuli of Experiment 4 in Tamaoka et al. (2005). For example, in the potential sentence (10a) the subject is marked by the dative case particle –*ni*, having a syntactic structure of [$_{s}$ NP-*ni* [$_{vp}$ NP-*ga* V]]. In this sentence, NP-*ni* is the subject whereas NP-*ga* is the object. Thus, case particles cannot provide the proper information to construct base structure. In contrast, according to case particle order suggesting that nominative proceeds dative and accusative, the canonical order should be (10b). If native Chinese speakers utilize case particles, they will have great difficulty processing the nominative-marked inanimate noun Greek-NOM. If they can understand Greek-NOM as actually being the object, and if they can comprehend that the dative-marked animate noun Takashi-DAT is the subject, then they can properly understand a potential sentence based on the base structure [$_{S}$ NP$_{\text{-subject (marked by the dative -}ni\text{)}}$ [$_{VP}$ NP$_{\text{-object (marked by the nominative -}ga\text{)}}$ V]].

(10) a. *Takasi ni* *girisyago ga* *kakerudarooka.*
Takashi DAT Greek NOM write-POTEN-wonder-Q
'Can Takashi write Greek?'

b. *girisyago ga* *Takasi ni* *kakerudarooka.*
Greek NOM Takashi DAT write-POTEN-wonder-Q

The processing of Japanese potential sentences by native Chinese speakers showed a trend that differed from non-potential active sentences. Tamaoka (2005, see Table 2 in Experiment 2) indicated that potential sentences in canonical order (M = 3,405 ms) did not significantly differ in reaction times from the same sentences in scrambled order (M = 3,774 ms). Taking null scrambling effects into account, it is possible to interpret that native Chinese speakers learning Japanese have not figured out the base structure of potential sentences, and therefore, the gap-filling parsing in the processing of potential sentences with scrambled order cannot apply to these Chinese speakers.

Before discussing the results of the response times, let's examine accuracies. Canonical order had an average of 69.1% with a high standard deviation of 23.9%, whereas the scrambled order had an average of 56.9% with an even higher standard deviation of 29.7%. This difference of 12.8% (Tamaoka 2005 showed 12.9%, but this was caused by a rounding error) between the canonical and scrambled orders was significant. Yet, the standard deviations of both the canonical and scrambled orders were very high, at over 20%. Individual participants are depicted in Figure 5, by plotting participants' (or students') accuracies on canonical order sentences on the horizontal axis and scrambled order on the vertical axis. To highlight participants' individual differences, the hierarchical cluster analysis for accuracies of canonical and scrambled orders revealed three clusters drawn on top of the plotting in Figure 5.

Let's consider three illuminating facts on individual differences of the clusters. First, three participants among the members of Cluster III in Figure 5, lying exactly on the horizontal axis, rejected all potential sentences with scrambled order as incorrect. In the scrambled order of OSV, an inanimate noun such as 'Greek' comes in the initial specifier position of the sentence, as in (10b), which is repeated below with the schematic structure.

(10) b. *girisyago ga* *Takasi ni* *kakerudarooka.*
Greek NOM Takashi DAT write-POTEN-wonder-Q
$[_{S}$ NP-$ga_{1}$ $[_{S'}$ NP-$ni$ $[_{VP}$ $gap_{1}$ V]]

'Greek' is in a subject position in (10b), marked by the nominative case particle –*ga*, which usually indicates the subject of a sentence. These three native Chinese speakers must have employed a simple and strict strategy that an inanimate subject

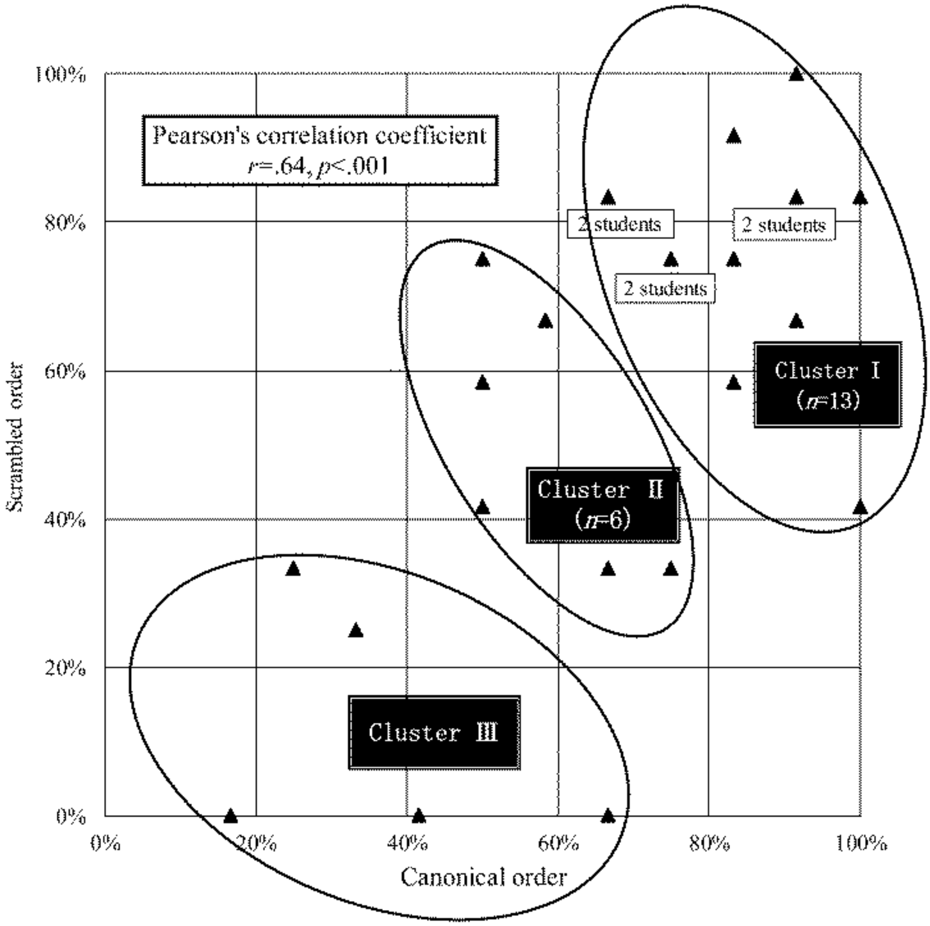



**Figure 5:** Accuracy plotting of canonical- and scrambled-ordered potential sentences (This figure is taken and translated from Tamaoka (2005: 103), *n* = 24.)

did not take the nominative case particle –*ga*, especially when placed in the initial specifier position, possibly indicating the subject of a sentence. For them, the following dative-marked animate noun adds the clear indication of incorrect marking.

As some researchers (e.g., Lamers and de Hoop 2005) suggest that animacy information plays a crucial role on language comprehension studies, native Chinese speakers may have utilized this strategy (see also Yamashita 2008). In fact, considering the typical events in daily life, animacy information is usually correct in that an animate actor as a subject acts upon an inanimate object, such as in 'My brother eats breakfast', 'My mother cooked clam chowder', and 'My father runs a vegetable shop'. The strategy of animacy, however, is not universally true, having some exceptions including potential sentences. The strategy of animacy with case particles must be deeply embedded in these three Chinese participants, allowing for no flexibility.

The second illuminating fact is that the five participants among the members of Cluster I in Figure 5 could process potential sentences at an accuracy rate higher than 80% in both canonical and scrambled orders. This is noteworthy in light of the possibility that a few native Chinese speakers could produce the base structure $[_{S}\ NP_{\text{-subject (marked by the dative -}\textit{ni}\text{)}}\ [_{VP}\ NP_{\text{-object (marked by the nominative -}\textit{ga}\text{)}}\ V]]$ for potential sentences, apparently moving beyond the conflicting nature of animacy and case particles. We must also bear in mind that these five participants were originally taken from a pool of 87 Chinese students majoring in Japanese language based on scores on a grammar test. This places them in approximately the top 5% of this group (more precisely 5.74%). It is quite possible that a few, possibly 5% of native Chinese speakers learning Japanese at a Chinese university, may understand potential sentences at a high rate of accuracy, which leaves open the great possibility that these learners could produce the base structure for potential sentences, and that they could also process scrambled-order potential sentences by gap-filling parsing.

The third fact is that accuracies on potential sentences of two participants among the members of Cluster II ranged between 40% and 60% in both canonical and scrambled order. They displayed a random pattern of decision making without a clear guideline for potential sentences. These two native Chinese speakers must have been puzzled to encounter potential sentences, in which animacy and case particles did not match correctly in terms of the nature of subject and object.

The contribution of individual differences to Japanese sentence processing must be measured as a reflection of Japanese language proficiency levels. This aspect was scrutinized by Tamaoka et al. (2010). They examined the degree of understanding of orally-presented single sentences in canonical and scrambled order based on Japanese ability. A listening comprehension test with a maximum of 8 points was conducted with 92 native Chinese speakers learning Japanese from one to three years at a university in Taiwan. Based on the test scores, a total of 48 participants were divided into higher (6–7 points), middle (4 points), and lower (2-1 points) listening comprehension groups (16 participants each) to undergo an experiment to investigate the understanding of orally-presented simple sentences (maximum of 11 points each). Two ditransitive active sentences in canonical and scrambled order were presented to the 48 native Chinese speaking participants. After canonical and scrambled active sentences were orally presented, the participants were asked two questions about the content of each sentence; one was related to the canonical sentence and another to the scrambled sentence. If a correct response was given, it was counted as one point. The study found a clear trend among the three groups. Scores of canonical ordered sentences significantly increased as comprehension levels increased: lower, middle, and higher. Scores of scrambled order sentences were comparatively lower for each group, with the higher group scoring significantly above the lower and middle groups, as in Table 15.

**Table 15:** Mean comprehension scores for canonical and scrambled sentences by group

| | Canonical | Scrambled |
|---|---|---|
| Lower | 6.19 | 5.06 |
| Middle | 7.44 | 4.88 |
| Higher | 8.31 | 7.19 |

Possible interpretations are that the lower group might have confused both canonical and scrambled order, which may have been caused by the difference in both parameter setting of the verb phrase and scrambling of subject and object noun phrases between Chinese and Japanese. The middle group was able to overcome the difference in word order of the verb phrase, and began to be able to handle the processing of Japanese sentences with canonical order. The higher group was able to establish a dependency between the initially-presented dative/accusative-marked phrase as *filler*, and its *gap* in the verb phrase (i.e., *filler-gap parsing*), resulting in higher scores in understanding both canonical and scrambled order sentences.

In sum, as seen in Figure 5, a great diversity was found among learners of Japanese with the same language background of Chinese. Once these individual differences are taken into consideration, it seems that researchers applying the theory of generative grammar to second language acquisition might be overly sensitive to the syntactic aspects of language. The shallow structure hypothesis by Clahsen and Felser (2006) cannot explain these individual differences in the manipulation ability of scrambled sentences. Rather, as White (2003) put forward, functional categories must be acquired at a relatively early stage of Japanese acquisition among native Chinese speakers. The progressive increase of scores in sentence comprehension shown by Tamaoka et al. (2010) must reflect the development of learners' facility with word order and advancement of parsing ability as native Chinese speakers improve their proficiency in the Japanese language.

# 5 Concluding remarks

Various studies with Chinese speaking learners of Japanese were reviewed in this chapter. As discussed, future studies can be categorized in three research areas: First, native English speakers showed an "awfully random" pattern of Japanese pitch accent acquisition regardless of the length of learning and proficiency (Taylor 2011a, 2011b, 2012), but native Chinese speakers displayed both random trends (Lee et al. 2006) and improvement as their learning progressed (Pen 2003). Since Chinese has tone accent, comparable to pitch accent, and since the position of the pitch accent

in each word is thoroughly taught when introducing Japanese vocabulary at universities in China (e.g., Hong 2010; Pan 2011; Zhang 2011; Zhao 2012; Zhou and Chen 2009, 2010, 2011a, 2011b), Chinese learners may exhibit some progress in acquiring Japanese pitch accent and advantage compared to learners of other L1 languages. Then, future studies on acquisition of pitch accent should pay special attention to dialectic influences in both Japanese and Chinese, differences in pitch accent patterns, and function of homophonic distinctions by controlling Japanese language proficiency of Chinese learners.

Second, due to the script similarity in kanji between L1 Chinese and L2 Japanese, Chinese learners demonstrate specific advantages and disadvantages in reading Japanese. Advantages are found in processing visually presented kanji compound words (e.g., Matsunaga 1999; Tamaoka 1997, 2000; Yamato and Tamaoka 2009, 2013). In contrast, because Chinese speakers heavily rely on their orthographic knowledge to understand Japanese words, their phonological processing of kanji compound words does not display advantage, and occasionally even showed inhibitory effects (e.g., H. Chiu 2003, 2003; Y. Chiu 2006; Hayakawa 2010; Hayakawa and Tamaoka 2012). In addition, semantic similarities and differences between L1 Chinese and L2 Japanese seem to exhibit complex processing trends in Chinese learners' understanding Japanese kanji compound words (e.g., Komori and Tamaoka 2010; Hayakawa and Tamaoka 2012). Therefore, future studies on advantages and disadvantages of L1 Chinese kanji knowledge for understanding L2 Japanese words should be conducted on the processing of phonologically similar/dissimilar words, kanji compound words with On- and Kun-readings, and semantic differences between the two languages, again with a population whose L2 Japanese language proficiency is controlled.

Third, Japanese and Chinese are considerably different in their syntactic features. Japanese word order is SOV while Chinese is SVO. Japanese has case particles while Chinese does not. Japanese allows scrambling (word permutation) while Chinese fundamentally does not. Because both languages are considered to have *longer linguistic distance* in syntax, difficulties in processing or understanding Japanese sentences are predicted (e.g., Horiba and Matsumoto 2008; Koda 1993, 2005). However, Chinese speakers' kanji knowledge allows them to allocate their Japanese learning hours to syntax, while those with no kanji language background spend many hours to memorize kanji. Thus they are likely to concentrate on syntax from the beginning stage of learning, resulting in high accuracies on morphosyntactic inflections of verbs (Chu et al. 2012). They also display the scrambling effect in processing SOV and OSV sentences (Tamaoka 2005; Tamaoka et al. 2010), as native Japanese speakers do (e.g., Aoshima, Phillips and Weinberg 2002; Koizumi and Tamaoka 2004, 2006; Mazuka, Itoh and Kondo 2002; Miyamoto and Takahashi 2002; Sakamoto 2002; Tamaoka 2004, 2006, 2010; Tamaoka et al. 2005). Difficulties by Chinese learners seem to come from slightly different verb and adjective usages

between L1 and L2, such as transitive/intransitive and active/passive. These usage differences should be investigated in future studies.

## References


Amano, Shigeaki and Tadahisa Kondo. 1999. *Nihongo no goi tokusei – Shinmitsudo [Lexical characteristics of Japanese words – Familiarity]*. Tokyo: Sanseido.

Amano, Shigeaki and Tadahisa Kondo. 2000. *Nihongo no goi tokusei – Shiyōhindo [Lexical characteristics of Japanese words – Frequency]*. Tokyo: Sanseido.

Andreev, Anton. 2002. Burugariajin gakushūsha no nihongo akusento – kēsu sutadī [Word-accent and Bulgarian learners of Japanese: A case study]. *Gengo kagaku ronshū [Journal of Linguistic Science, Tohoku University]* 6. 1–12.

Aoshima, Sachiko, Colin Phillips and Amy Weinberg. 2002. Active filler effects and reanalysis: A study of Japanese *Wh*-scrambling constructions. *University of Maryland Working Papers in Linguistics* 12. 1–24.

Ayusawa, Takako. 1998. Nihongo gakushūsha ni totte no Tōkyō akusento [The Tokyo accent for Japanese learners]. *Gekkan Gengo [Monthly Linguistics]* 27(1). 70–75.

Bates, Elizabeth and Brian MacWhinney. 1987. Competition, variation, and language learning. In Brian MacWhinney (ed.), *Mechanisms of language acquisition*, 157–193. Hillsdale, NJ: Erlbaum.

Bunka-chō [Agency for Cultural Affairs in Japan]. 1978. *Chūgokugo to taiō suru kango [Kanji-compound words corresponding to Chinese]*. Tokyo: Agency for Cultural Affairs in Japan.

Bunka-chō [Agency for Cultural Affairs in Japan]. 2011. *Kokunai no nihongo kyōiku no gaiyō [The outline of Japanese language education in Japan]*. Tokyo: Agency for Cultural Affairs in Japan.

Cai, Feng-Xiang and Norio Matsumi. 2009. Chūgokugo o bogo to suru jōkyū nihongo gakushūsha ni okeru nihongo kanji tango no shori katei – dōkongo to hidōkongo o mochiita gengokan puraimingu-hō ni yoru kentō [The processing of kanji words in Chinese proficient learners of Japanese: Using cross-language priming paradigm with cognates and non-cognates]. *Nihongo kyōiku* 141. 13–24.

Chang, Franklin. 2015. The role of learning in theories of English and Japanese sentence processing. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Chen, Yu Min. 2002. Nihongo niji kanji goi to sore ni taiō suru chūgokugo niji kanji goi wa onazi ka – Taiwan oyobi chūgoku no chūgokugo to no hikaku [Are two-kanji compounds in Japanese the same as those in Chinese – A comparison with the Chinese language in Taiwan and the mainland China]. *Gengo Bunka to Nihongo Kyōiku [Language Culture and Japanese Education]* 24. 40–53.

Chiu, Hsueh Chin. 2002. Taiwanjin nihongo gakushūsha ni okeru nihongo kanji jukugo no shori katei – Nitchū 2-gengo kan no dōkongo to hi-dōkongo no hikaku [The processing routes of Japanese kanji words for the Taiwanese learners of Japanese: The comparison of cognates and non-cognates between Japanese and Chinese]. *Hiroshima Daigaku Daigakuin Kyōikugaku Kenkyūka Kiyō – Dai 2-bu [Bulletin of Graduate Studies of Education at Hiroshima University – Part 2]* 51. 357–365.

Chiu, Hsueh Chin. 2003. Taiwanjin nihongo gakushūsha no nihongo kanji jukugo no on'in shori nitsuite – tango taipu/tango no shūtoku nenrei/shūzyukudo no kanten kara no kōsatsu [The phonological processing of Japanese kanji words by Taiwanese learners of Japanese: An analysis of word type, age of acquisition, and fluency level]. *Nihongo Kyōiku [Journal of Japanese Language Teaching]* 116. 89–98.

Chiu, Yu Yuan. 2006. Taiwanjin nihongo gakushūsha ni okeru nihongo tango no chōkakuteki ninchi – Nihongobun no imihandan kadai o mochiita kentō [Auditory recognition of Japanese words in Taiwanese learners of Japanese: Discussion on meaning decision task of Japanese sentences]. *Hiroshima Daigaku Daigakuin Kyōikugaku Kenkyūka Kiyō – Dai 2-bu [Bulletin of Graduate Studies of Education at Hiroshima University – Part 2]* 55. 267–273.

Chiu, Yu Yuan. 2007. Taiwanjin nihongo gakushūsha ni okeru nihongo tango no chōkakuteki ninchi – Dōkongo, hi-dōkongo, hiragana-tango, katakana-tango no hikaku [Auditory recognition of Japanese words in Taiwanese learners of Japanese: Comparison of Chinese-Japanese cognates with non-cognates, hiragana words, and katakana words]. *Nihongo Kyōiku [Journal of Japanese Language Teaching]* 132. 108–117.

Chu, Xiang Juan, Katsuo Tamaoka and Yuko Yamato. 2012. Shokyū chūgokujin nihongo gakushūsha no *te*-kei shūtoku [Acquisition of *te*-form by native Chinese speakers learning Japanese at the beginner's level]. *Nihon Kyōka Kyōiku Gakkaishi [The Bulletin of Japanese Curriculum Research and Development]*. 35(2). 63–72.

Clahsen, Harald and Claudia Felser. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27. 3–42.

Costa, Albert, Alfonso Caramazza and Núria Sebastián-Gallés. 2000. The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(5). 1283–1296.

Cutler, Anne and Takashi Ōtake. 1999. Pitch accent in spoken-word recognition in Japanese. *Journal of the Acoustical Society of America* 105. 1877–1888.

de Groot, Annette M. B., Philip Delmaar and Stephen J. Lupker. 2000. The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *Quarterly Journal of Experimental Psychology* 53A. 397–428.

Dijkstra, Tom and Walter J. B. van Heuven. 2002. The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* 5. 175–197.

Djojomihardjo, Muljani, Keiko Koda and Danny R. Moates. 1994. Development of L2 word recognition. In Qicheng Jing, Houcan Zhang and Danling Peng (eds.), *Information processing of Chinese language*. 153–161. Beijing: Beijing Normal University Publishing.

Dryer, Matthew S. 2012. Order of subject, object and verb. The world atlas of language structures online. In Matthew S. Dryer and Martin Haspelmath (eds.), Munich: Max Planck Digital Library. [Online] Retrieved on 21 October 2012 from: http://wals.info/chapter/81.

Endo, Shotoku. 1986. *Hayawakari chūgoku kantaigi [Quickly understanding of Chinese character simplification]*. Tokyo: Kokusho Kankokai.

Fan, Jing and Shao Hua Wu. 2006. Xi'an waiguoyu xueyuan 2002 ji 2003 nian riyu zhuanye kaoshi jieguo fenxi [Results of the Japanese specialized test at the Xi'an International Studies University in 2002 and 2003]. *Shanxi shifan daxue xuebao (Zhexue shehui kexueban) [Journal of Shanxi Normal University (Philosophy and Social Sciences Edition)]* 35. 415–419.

Fukuoka, Masako. 2008. Kankokujin nihongo gakushūsha no akusento shūtoku ni okeru bogo kanshō: Gotō haretsuon o fukumu go no akusento [Interference of the mother tongue on acquisition of Japanese accent by native Korean speakers learning Japanese: An accent of words including an explosive sound at word initial position], *Mie daigaku kokusai kōryū sentā kiyō [Bulletin of the Center for International Education and Research at Mie University]* 3. 45–59.

Green, David. W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1. 67–81.

Hadamitzky, Wolfgang and Mark Spahn. 1981. *Kanji and kana: A handbook and dictionary of the Japanese writing system*. Tokyo: Charles E. Tuttle Company.

Hattori, Shiro. 1951. *Onsei-gaku [Phonology]*. Tokyo: Iwanami Shoten.

Hawkins, Roger. 2001. *Second language syntax: A generative introduction.* Malden, MA: Blackwell Publishing.

Hayakawa, Kyoko. 2010. Chūgokugo o bogo to suru nihongo gakushūsha ni yoru kanjigo no on'in shori [Phonological processing of kanji compounds by native Chinese speakers learning Japanese]. *Nitchū Gengo Kenkyū to Nihongo Kyōiku [Journal of Japanese and Chinese Linguistics and Japanese Language Education]*. 3. 100–110.

Hayakawa, Kyoko and Katsuo Tamaoka. 2012. Chūgokujin/Kankokujin nihongo gakushūsha ni yoru chōkaku/shikaku-teiji no gengokan-dōkeidōgi/gengokan-ikeidōgi no niji kanjigo no shori [Processing of auditory- and visually-presented orthographically/semantically similar and different two-kanji compound words by native Chinese and Korean speakers learning Japanese]. *Koide kinen nihongo kyōiku kenkyūkai ronbunshū [Koide Memorial Journal of Japanese Language Education]*. 20. 17–32

Hayata, Teruhiro. 1999. *Onchō no taiporojī [Typology of tones]*. Tokyo: Taishukan Shoten.

He, Yuan Jian. 2000. Di 7 zhang: Yiwenci yiwei yu hanyu yiwenci de yufa fenbu [Chapter 7: *Wh*-movements and phrasal rules of Chinese questions]. In Yang Shen, Yuanjian He, and Yang Gu (eds.), *Shengcheng yufa lilun yu hanyu yufa yanjiu [The theory of generative grammar and the study of Chinese grammar]*, 159–190. Haerbin, China: Heilongjiang Jiaoyu Chubanshe.

Hirayama, Teruo. 1957. *Nihongo Onchō no kenkyū [Study on Japanese accents]*. Tokyo: Meiji Shoin.

Hirayama, Teruo. 1968. *Nihongo no hōgen [Dialects in Japan]*. Tokyo: Kodansha.

Hirose, Hitoshi. 1998. Identifying the On- and Kun-readings of Chinese characters: Identification of On versus Kun as a strategy-based judgment. In Chen Kan Leong and Katsuo Tamaoka (eds.), *Cognitive processing of the Chinese and the Japanese languages*, 375–394. Dordrecht, Netherlands: Kluwer Academic Publishers.

Hishinuma, Toru. 1983. Nihongo to chūgokugo no jōyō jitai [commonly-used kanji in Japanese and Chinese]. *Chūgoku Kenkyū Geppō [Monthly-journal of Chinese Studies]* 428. 1–20.

Hishinuma, Toru. 1984. chūgokugo no hyōjun jitai to nihon no jōyō jitai [The standard orthography in Chinese and the common orthography in Japanese]. *Nihongogaku [Japanese Linguistics]* 3. 32–40.

Hong, Lin. 2010. *Jichu riyu zonghe jiaocheng 1 [Basic Japanese general course 1]*. Beijing, China: Gaodeng Jiaoyu Chubanshe.

Hong, Shu Juan. 2004. Kikitori no purosesu ni okeru kanji no eikyō – Taiwanjin nihongo gakushūsha (dikuteishon bunseki) o chūshin ni [Effects of kanji for listening process – focusing on Taiwanese learning Japanese (dictation analysis)]. *Tsukuba Daigaku Chiiki Kenkyū [Area Studies Tsukuba]* 22. 155–176.

Horiba, Yukie and Junko Matsumoto. 2008. Bunmyakubun no fukuzatsusa ga bunpō kōmoku no rikai ni oyobosu eikyō: Bogo haikei no kotonaru dai-2 gengo gakushūsha no hikaku kara [Effects of syntactic complexity on comprehension of sentences with contexts: From a comparative study of the second language learners with different backgrounds of mother tongues]. *Scientific approaches to languages (Kanda gaigo daigaku gengo kenkyū sentā kiyō [The Bulletin of the Center for Language Sciences, Kanda University of International Studies])* 7. 181–207.

Hu, Marrcella Hsuch-chao and Nation I. S. Paul. 2000. Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language* 13. 403–430.

Huang, C-T. James. 1981. Move *wh* in a language without *wh*-movement. *Linguistic Review* 82(1). 369–416.

Ishida, Toshiko. 1986. Eigo chūgokugo, kankokugo kenbetsu nihongoryoku no bunseki [Analysis of Japanese proficiency among areas of English, Chinese and Korean]. *Nihongo Kyōiku [Journal of Japanese Language Teaching]* 58. 162–194.

Iwata, Rei. 2001. Chūgokugo no seichō to akusento [Chinese tones and accents]. *Onsei Kenkyū [Journal of the Phonetic Society of Japan]* 5(1). 18–27.

Japan Foundation [Kokusai Kōryū Kikin] and Japan Educational Exchange and Services [Nihon Kokusai Kyōiku Shien Kyōkai]. 2002. *Nihongo nōryoku shiken shutsudai kijun – kaiteiban [The production standard of Japanese proficiency tests – Revised version]*. Tokyo: Japan Foundation and Japan Educational Exchange and Services.

Japan Foundation [Kokusai Kōryū Kikin]. 2011. *Kaigai no nihongo kyōiku no genjō: Nihon-go kyōiku kikan chōsa 2009 – Gaiyō [Demographic figures of Japanese language education outside of Japan: A survey on institutes of Japanese language education in 2009 – Summary]*. Tokyo: Japan Foundation.

Kaiho, Hiroyuki and Yukimura Nomura. 1983. *Kanji jōhō shori no shinrigaku [Psychology of kanji information processing]*. Tokyo: Kyoiku Shuppan.

Kayamoto, Yuriko. 1995a. *Nihongo kanji no yomi ni oyobosu bogo no on'in jōhō no eikyō – chūgokugo bogo washa no baai [Effects of phonological information from mother-tongue in pronouncing Japanese kanji – A case study of native Chinese speakers]*. Hiroshima: Hiroshima University MA thesis.

Kayamoto, Yuriko. 1995b. Dōitsu kanji ni okeru chūgokugoon to nihongo no on'yomi no ruijido ni kansuru chōsa [Similarities and differences between readings of Chinese characters and On-readings of Japanese kanji]. *Hiroshima Daigaku Nihongo Kyōikugakka Kiyō [Bulletin of Graduate School of the Japanese Language Education at Hiroshima University]* 5. 67–75.

Kayamoto, Yuriko. 1996. Nihongo kanji to chūgokugo kanji no keitaiteki/on'inteki sai ga chūgokugo bogowasha ni yoru nihongo kanji no yomi ni oyobosu eikyō [Effects of orthographic/semantic differences between Chinese characters and Japanese kanji on naming Japanese kanji by native Chinese speakers]. *Hiroshima Daigaku Kyōiku Gakubu Kiyō [Bulletin of Department of Education at Hiroshima University]* 45. 345–352.

Kayamoto, Yuriko. 2000. Nihongo o gakushūsuru chūgokugo bogo washa no kanji no ninchi [Processing phonological information: Recognition of Japanese characters by advanced- and superior-level native speakers of Chinese]. *Kyōiku Shinrigaku Kenkyū [Journal of Educational Psychological Research]* 48. 315–322.

Kayamoto, Yuriko. 2002. Goi handan kadai to meimei kadai ni okeru chūgokugo bogo washa no nihongo kanji akusesu [Lexical access to Japanese kanji by native speakers of Chinese: Evidence from lexical decision and naming task]. *Kyōiku Shinrigaku Kenkyū (Journal of Educational Psychology)*. 50. 436–445.

Kess, Joseph F. and Tadao Miyamoto 1999. *The Japanese mental lexicon: Psycholinguistic studies of kana and kanji processing*. Philadelphia, PA: John Benjamins Publishing Company.

Kim, Mi-Ryoung and San Duanmu. 2004. "Tense" and "lax" stops in Korean. *Journal of East Asian Linguistics* 13. 59–104.

Kindaichi, Haruhiko. 1974. *Kokugo akusento no shiteki kenkyū – Genri to hōhō [Historical study of Japanese accent – Principle and method]*. Tokyo: Hanawa Shobo.

Kitahara, Mafuyu. 2006. Akusento tairitsu no bunpu ni tsuite [On the distribution of accentual oppositions]. *Bunpō to onsei [Speech and Grammar]* 5. 147–158.

Koda, Keiko. 1993. Transferred L1 strategies and L2 syntactic structure in L2 sentence comprehension. *The Modern Language Journal* 77. 490–500.

Koda, Keiko. 2005. *Insights into second language reading: A cross-linguistic approach*. New York: Cambridge University Press.

Koizumi, Masatoshi. 2015. Experimental syntax: word order in sentence processing. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Koizumi, Masatoshi and Katsuo Tamaoka. 2004. Cognitive processing of Japanese sentences with ditransitive verbs. *Gengo Kenkyū* 125. 173–190.

Koizumi, Masatoshi and Katsuo Tamaoka. 2006. Bun kaiseki jikken ni yoru nihongo fukusirui no gojun no hantei [Determination of basic word order of adverbs in Japanese by a sentence-processing experiment]. *Ninchi Kagaku [Cognitive Studies]* 13. 392–403.

Koizumi, Masatoshi and Katsuo Tamaoka. 2010. Psycholinguistic evidence for the VP-internal subject position in Japanese. *Linguistic Inquiry* 41. 663–680.

Kokuritsu Kokugo Kenkyūjo [National Institute for Japanese Language and Linguistics]. (1964). *Gendai zasshi kyūjūsshu no yōgo yōrei dai-3 bunsatsu bunseki [Usages and examples in 90 different modern journals: The 3rd separate volume – analysis]*. Tokyo: Shuei Shuppan.

Komori, Kazuko. 2005. Daini gengo toshite no nihongo no bunshōrikai ni okeru daiichi gengo no tango ninchi shori hōryaku no ten'i – shikaku nyūryoku to chōkaku nyūryoku no sōi o chūshin ni [Transfer of a word recognition processing strategy of the first language on text understanding of Japanese as a second language – Focusing on differences between visual input and auditory input]. *Yokohama kokuritsu daigaku ryūgakusei sentā kiyō [Journal of the International Student Center]* 12. 17–39.

Komori, Kazuko and Katsuo Tamaoka. 2010. Chūgokujin nihongo gakushūsha niyoru dōkei ruigigo no ninchi shori [The cognitive processing of lexical homographs by native Chinese speakers learning Japanese]. *Lexicon Forum*, 5. 165–200.

Komori, Kazuko, Junko Mikuni and Atsuko Kondo. 2004. Bunshō rikai o sokushin suru goi chishiki no ryōteki sokumen – kichigo ritsu no ikichi tansaku no kokoromi [What percentage of known words in a text facilities reading comprehension?: A case study for exploration of the threshold of known words coverage]. *Nihongo Kyōiku [Journal of Japanese Language Teaching]* 120. 83–92.

Kubozono, Haruo and Satoshi Ota. 1998. *On'in kōzō to akusento [Phonological structure and accents]*. Tokyo: Kenkyusha.

Lamers, Monique and Helen de Hoop. 2005. Animacy information in human sentence processing: An incremental optimization of interpretation approach. Henning Christiansen, Peter Rossen Skadhauge and Jørgen Villadsen (eds.), *Constraint solving and language processing – Lecture notes in computer science volume 3438*, 158–171, Berlin: Springer.

Lee, Wood Hung, Ken'ichiro Murashima and Yasuhiro Shirai. 2006. Prosodic development in the acquisition of Japanese: A longitudinal study of three native speakers of Chinese. *Nihongakukan [Journal of Japan Studies]* 10. 38–51.

Leong, Che Kan. 1986. What does accessing a morphemic script tell us about reading and reading disorders in an alphabetic script? *Annals of Dyslexia* 36. 82–102.

Lin, Jo-Wang. 1998. On existential polarity *wh*-phrases in Chinese. *Journal of East Asian Linguistics* 7. 219–255.

Liu, Jia Qi. 2010. Chūgokugo bogo washa (pekin to shanhai shusshinsha) ni yoru fukugō dōsi no Tōkyō akusento no shūtoku [A study on the acquisition of Japanese accent of compound verbs by Chinese speaking learners (from Beijing and Shanghai)]. *Waseda Nihongo Kyōikugaku [the Study of Japanese Education in Waseda]* 8. 15–28.

Matsunaga, Sachiko. 1999. The role of kanji knowledge transfer in acquisition of Japanese as a foreign language. *Japanese-Language Education around the Globe*, 9. 87–100.

Mazuka, Reiko, Kenji Itoh and Tadahisa Kondo. 2002. Costs of scrambling in Japanese sentence processing. In Mineharu Nakayama (ed.), *Sentence processing in East Asian languages*, 131–166. Stanford, CA: CSLI Publications.

Miller, Roy Andrew. 1967. *The Japanese language*. Chicago, IL: The University of Chicago Press.

Miyamoto, Edson T. and Shoichi Takahashi. 2002. Sources of difficulty in the processing of scrambling in Japanese. In Mineharu Nakayama (ed.), *Sentence processing in East Asian languages*, 167–188. Stanford, CA: CSLI Publications.

Miyaoka Yayoi, Katsuo Tamaoka and Hiromu Sakai. 2011. Nihongo goi tesuto no kaihatsu to shinraisei: Chūgokugo o bogo tosuru nihongo gakushūsha no dēta ni yoru tesuto hyōka [Development of reliability of Japanese lexical knowledge test: Test evaluation by data taken from native Chinese speakers learning Japanese]. *Hiroshima Keizai Daigaku Kenkyū Ronshū [Bulletin of Hiroshima University of Economic]* 34(1). 1–18.

Morrison, Catriona M. and Andrew W. Ellis. 1995. Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition* 21. 116–133.

Nagatomo, Kazuhiko. 1997. Dōsi *te*-kei ni kakawaru on'in kisoku no shūtoku to gengo no fuhensei [Acquisition of phonological rules involved in the verbal *te*-form]. *Daini gengo to shite no nihongo no shūtoku kenkyū [Acquisition of Japanese as a Second Language]* 1. 1–8.

Nakato, Yasue. 2001. Tango yomiage ni okeru kankokujin nihongo gakushūsha no pitchi jitsugen [Pitch realization of sounding words by Korean learners of Japanese]. *Nihongo kyōiku [Japanese Language Education]* 109. 80–89.

Nation, I. S. Paul. 2001. *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

Nishinuma, Yukihiro. 1994. How do the French perceive tonal accent in Japanese? Experimental evidence. *Proceedings of the fourth international conference on spoken language processing (ICSLP)*. Philadelphia, USA, 646–649.

Otake, Takashi and Anne Cutler. 1999. Perception of suprasegmental structure in a nonnative dialect. *Journal of Phonetics* 27. 229–253.

Otake, Takashi. 2002. A role of prosody in spoken-word recognition. *Journal of the Phonetics Society of Japan* 6(2). 56–65.

Pan, Shou-Jun. 2011. *Jichu riyu zonghe jiaocheng 2 [Basic Japanese general course 2]*. Beijing, China: Gaodeng Jiaoyu Chubanshe.

Pan, Xin Ying. 2003. Taiwanjin no nihongo akusento chikaku ni okeru shoyōin [Factors in Japanese accent perception by Taiwanese]. *Tsukuba ōyō gengogaku kenkyū [Study of Applied Linguistics in Tsukuba]* 10. 83–96.

Saito, Sumio. 2006. *Nihongo onseigaku nyuūmon – kaiteiban [Introduction to Japanese phonology – Revised version].* Tokyo: Sanseido.

Sakamoto, Tadashi. 1993. Eigo washa ni okeru '*te*-kei' keiseikisoku no shūtoku ni tsuite [Acquisition of *te*-form constructing rules by native English speakers]. *Nihongo Kyōiku [Journal of Japanese Language Teaching]* 80. 125–135.

Sakamoto, Tsutomu. 2002. Processing filler-gap constructions in Japanese. In Mineharu Nakayama (ed.), *Sentence processing in East Asian languages*, 189–221. Stanford, CA: CSLI Publications.

Sawasaki, Koichi. 2006. *L2 reading by learners of Japanese: A comparison of different L1s.* Columbus, OH: The Ohio State University dissertation.

Sekiguchi, Takahiro. 2006. Effects of lexical prosody and word familiarity on lexical access of spoken Japanese words. *Journal of Psycholinguistic Research* 35. 369–384.

Shibata, Takeshi and Ritei Shibata. 1990. Akusento wa dōongo o donoteido benbetsu shiuru no ka?: Nihongo, eigo, chūgokugo no baai [How much can accent distinguish homophones?: cases of Japanese, English and Chinese]. *Keiryō Kokugogaku [Mathematical Linguistics]* 17. 317–327.

Shirai, Yasuhiro. 2015. L2 acquisition of Japanese. In Mineharu Nakayama (ed.), *Handbook of Japanese psycholinguistics*. Boston: De Gruyter Mouton.

Stahl, Steven A. and William E. Nagy 2006. *Teaching word meanings*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sugito, Miyoko. 1982. *Nihongo no akusento kenkyū [The study on Japanese accents].* Tokyo: Sanseido.

Sugito, Miyoko. 1989. *Hana to hana [Nose and flower].* Osaka, Japan: Izumi Shoin.

Sugito, Miyoko. 2006. *Nihongo onseigaku nyūmon – kaiteiban [Introduction to Japanese phonology – Revised version].* Tokyo: Sanseido.

Sugito, Miyoko. 2012. *Nihongo no akusento, eigo no akusento [Japanese accent and English accent].* Tokyo: Hituzi Syobo.

Sukegawa, Yasuhiko. 1999. Burajirujin nihongo gakushūsha no 2-mōra to 3-mōra go no pitchi jitsugen [Pitch realization of 2/3-mora-words by Brazilian learners of Japanese]. *Onsei Kenkyū [Journal of the Phonetic Society of Japan]* 3(3). 13–35.

Tamaoka, Katsuo. 1991. Psycholinguistic nature of the Japanese orthography. *Gengo Bunka Kenkyū [Studies in Language and Literature]* (Matsuyama University). 11(1). 49–82.

Tamaoka, Katsuo. 1997. Chūgokugo to eigo o bogo tosuru nihongo gakushūsha no kanji oyobi kana hyōki goi no shori hōryaku [The processing strategy of words presented in kanji and kana by Chinese and English speakers learning Japanese]. *Gengo Bunka Kenkyū [Studies in Language and Literature]* (Matsuyama University) 17(1). 65–77.

Tamaoka, Katsuo. 2000. Chūgokugo oyobi eigokei nihongo gakushūsha no bogo no hyōki keitai ga nihongo no on'inshori ni oyobosu eikyō [The effects of mother-tongue scripts on the phonological processing of the Japanese language by Chinese and English speakers]. *Dokusho Kagaku [The Science of Reading]* 44(3). 83–94.

Tamaoka, Katsuo. 2005. Chūgokugo o bogo tosuru nihongo gakushūsha niyoru seijun/kakimaze gojun no nōdōbun to kanōbun no rikai [Comprehension of Japanese active and potential sentences with canonical and scrambled word orders by native Chinese speakers learning the Japanese language]. *Nihongo Bunpō [Japanese Grammar]* 5(2). 92–109.

Tamaoka, Katsuo, Kim Kirsner, Yushi Yanase, Yayoi Miyaoka and Masahiro Kawakami. 2002. A web-accessible database of characteristics of the 1,945 basic Japanese kanji. *Behavior Research Methods, Instruments, & Computers* 34. 260–275.

Tamaoka, Katsuo, Hiromu Sakai, Jun Kawahara, Yayoi Miyaoka, Hyun Jung Lim and Masatoshi Koizumi. 2005. Priority information used for the processing of Japanese sentences: Thematic roles, case particles or grammatical functions? *Journal of Psycholinguistic Research* 34. 273–324.

Tamaoka, Katsuo and Marcus Taft. 2010. The sensitivity of native Japanese speakers to On and Kun kanji readings. *Reading and Writing* 23. 957–968.

Tamaoka, Katsuo, Hsueh Chin Chiu, Yayoi Miyaoka and Sachiko Kiyama. 2010. Chūgokugo o bogo to suru nihongo gakushūsha ni yoru kakimaze gojun no bunrikai: Chōkai nōryoku de waketa jōi, chūi, kai gun no hikaku [Understanding of sentences with scrambled word order by native Chinese speakers learning Japanese: Comparisons of higher, middle, and lower groups divided by listening comprehension]. *Nihongo Bunpō [Japanese Grammar]* 5(2). 92–109.

Tamaoka, Katsuo and Sachiko Kiyama. 2013. The effects of visual complexity for Japanese kanji processing with high and low frequencies. *Reading and Writing* 26. 205–223.

Trenkic, Danijela. 2002. Form-meaning connections in the acquisition of English articles. *The EUROSLA (European Second Language Association) Yearbook* 2. 115–133.

Taylor, Becky. 2011a. Variability and systematicity in individual learners' Japanese accent. *Pozanan Studies in Contemporary Linguistics* 47. 146–158.

Taylor, Becky. 2011b. Do English learners of Japanese produce isolated nouns with standard Japanese lexical accent? *Second Language* 10. 15–31.

Taylor, Becky. 2012. *Eigo washa ni yoru nihongo no go-akusento no shūtoku [The acquisition of Japanese lexical accent by English speakers]*. Nagoya, Japan: Nagoya University dissertation.

Toda, Takako. 1999. Nihongo gakushūsha ni yoru gairaigo shiyō no jittai to akusento shūtoku ni kansuru kōsatu – Eigo, chūgokugo kankokugo washa no kaiwa dēta ni motozuite [Usage and acquisition of accent of English loanwords by Japanese learners: An analysis based on conversational data from English, Chinese and Korean Speakers]. *Bungei gengo kenkyū – Gengo-hen [The Journal of Literature and Language Studies – Language]* 36. 89–111.

Vance, Timothy J. 2008. *The sounds of Japanese.* Cambridge, UK: Cambridge University Press.

van Heuven, Walter. J. B., Tom Dijkstra and Jonathan Grainger. 1998. Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language* 39. 458–483.

van Heuven, Walter J. B., Herbert Schriefers, Tom Dijkstra and Peter Hagoort. 2008. Language conflict in the bilingual brain. *Cerebral Cortex* 18. 2706–2716.

Verdonschot, Rinus G., Wido La Heij, Katsuo Tamaoka, Sachiko Kiyama, Wen Ping You and Niels O. Schiller. 2013. The multiple pronunciations of Japanese kanji: a masked priming investigation. *Quarterly Journal of Experimental Psychology.* On-line first.

White, Lydia. 2003. *Second language acquisition and Universal Grammar.* Cambridge UK: Cambridge University Press.

Xu, Bao Hua and Zhen Zhu Tang. 1988. *Shanghaishiqu fangyan zhi [A study on the Shanghai dialect].* Shanghai, China: Shanghai Jiaoyu Chubanshe [Shanghai Educational Publishing].

Yamashita, Hiroko. 2008. Effects of sentence processing strategy proximity on the comprehension of second languages. *Second Language* 7. 43–82.

Yamato, Yuko and Katsuo Tamaoka. 2009. Chūgokujin nihongo gakushūsha no nihongo kanjigo no shori ni okeru bogo no eikyō [Effects of mother tongue by native Chinese speakers learning Japanese on processing kanji-compound words]. *Kotoba no kagaku [The Science of Languages]* 22. 117–135.

Yamato, Yuko and Katsuo Tamaoka. 2013. Chūgokugo bogowasha to kankokugo bogowasha no nihongo tekisuto no yomi shori ni okeru bogo to no gengoteki ruijisei no eikyō [Effects of linguistic similarity on Japanese text processing by native Chinese and Korean speakers]. *Koide kinen nihongo kyōiku kenkyūkai ronbunshū [Koide Memorial Journal of Japanese Language Education]* 21. 229–268.

Yang, Wen Jin. 2011. Taiwanjin nihongo gakushūsha no akusento: Akusento no jittai/shūtoku ni kakawaru yōin kyōiku e no yōbō [Accents of Taiwanese learning Japanese: Situation of accent acquisition, its factors and educational implement]. *Hōgen onsei kenkyū [Dialect and Phonetic Studies]* 5. 1–16.

Yin, Song. 2002. Patān gakushū wa rikai o sokushin saseruka – rajio nyūsu no chōkai no baai [Does pattern teaching improve students' listening comprehension?: The case of radio news comprehension]. *Nihongo Kyōiku [Journal of Japanese Language Teaching]* 112. 35–44.

Yokosawa, Kazuhiko and Michio Umeda. 1988. Processes in human Kanji-word recognition. *Proceedings of the 1988 IEEE International Conference on Systems, Man, and Cybernetics* 377–380. August 8–12, Beijing and Shenyang, China.

You, Ru Jie. 2004. *Hanyu fangyanxue jiaocheng [A study on Chinese dialects].* Shanghai, China: Shanghai Jiaoyu Chubanshe [Shanghai Educational Publishing].

Zevin, Jason D. and Mark S. Seidenberg. 2002. *Age of acquisition effects in word reading and other tasks. Journal of Memory and Language* 47. 1–29.

Zhang, Hui-Fen. 2011. *Jichu riyu zonghe jiaocheng 3 [Basic Japanese general course 3].* Beijing, China: Gaodeng Jiaoyu Chubanshe.

Zhao, Hua-Min. 2012. *Jichu riyu zonghe jiaocheng 4 [Basic Japanese general course 4].* Beijing, China: Gaodeng Jiaoyu Chubanshe.

Zhou, Ping and Xiao-Fen Chen. 2009. *Xin bian riyu 1 [A newly-edited Japanese language 1].* Shanghai, China: Shanghai Waiyu Jiaoyu Chubanshe.

Zhou, Ping and Xiao-Fen Chen. 2010. *Xin bian riyu 2 [A newly-edited Japanese language 2].* Shanghai, China: Shanghai Waiyu Jiaoyu Chubanshe.

Zhou, Ping and Xiao-Fen Chen. 2011a. *Xin bian riyu 3 [A newly-edited Japanese language 3].* Shanghai, China: Shanghai Waiyu Jiaoyu Chubanshe.

Zhou, Ping and Xiao-Fen Chen. 2011b. *Xin bian riyu 4 [A newly-edited Japanese language 4].* Shanghai, China: Shanghai Waiyu Jiaoyu Chubanshe.

# Subject index